Corr
56

RATIO

edited in co-operation with
H. B. Acton (London), A. Church (Princeton),
H. Cherniss (Princeton), G. Henry-Hermann (Bremen),
K. R. Popper (London) and J. W. N. Watkins (London)
by
Julius Kraft (Frankfurt a. M.)

VOLUME I
1957-8

BASIL BLACKWELL
OXFORD

CONTENTS

Volume I

Number 1, December 1957

	Page
Philosophical observations on Mathematical Logic and on investigations into the Foundations of Mathematics, by W. ACKERMANN	1
The Aim of Science, by K. R. POPPER	24
Religion, Experience and Metaphysics, by J. KRAFT	36
Conscience and Reason, by G. HECKMANN	52
A Samson in the Temple of Philosophy, by J. W. N. WATKINS	68
On the Operational Foundations of Logic and Mathematics, by G. H. MÜLLER	85

Number 2, December 1958

On the Status of Science and of Metaphysics, by K. R. POPPER	97
The Philosophy of R. G. Collingwood, by E. W. F. TOMLIN	116
The Marxist-Leninist Theory of Religion, by H. B. ACTON	136
On Leśniewski's Ontology, by C. LEJEWSKI	150
REVIEWS:	
Leonard Nelson: *System of Ethics* (Brand Blanshard)	177
Ernst Topitsch: *Vom Ursprung und Ende der Metaphysik* (H. Kaestner)	182
Kurt Reidemeister: *Raum und Zahl* (W. Ackermann)	185

and make proper important volumes delay return of Building Thanks

PHILOSOPHICAL OBSERVATIONS ON MATHEMATICAL LOGIC AND ON INVESTIGATIONS INTO THE FOUNDATIONS OF MATHEMATICS

By Wilhelm Ackermann

In conversations with acquaintances interested in philosophy I have observed that one is here often dealing with the representatives of two groups. The one group reject logico-mathematical foundational research on the most varied grounds and deny its philosophical significance. In so far as this attitude rests on objections which must be taken seriously we shall consider it below; in so far as it rests on an aversion to everything mathematical no discussion is possible, not even on the philosophical significance of mathematics. The other group consists of people who are in general interested in mathematics, but to whom the results of modern foundational research are unknown and who have only a knowledge of mathematics from its principal side, corresponding to the state of knowledge of fifty and more years ago. This is the more deplorable because just in the last decades an impetuous development has taken place in the field of mathematical logic and mathematical foundational research, with results that cannot be disregarded.

Consequently I shall deal in what follows with some of the main objections to mathematical logic and attempt at the same time to expound some of its most important results. These results are familiar to the student of foundations; this exposition is addressed to another quarter, as can be seen from the informal character of the statements which follow. No systematic and comprehensive presentation of the subject is given here; we are concerned with isolated observations as they have arisen from conversation. They are the observations of a mathematician rather than a philosopher, but of a mathematician who is interested in philosophy.

First a warning! It should be clearly understood that mathematical logic as true logic has nothing whatever to do with any philosophical view of the theory of knowledge. In spite of this one is always hearing the view that to be concerned with mathematical logic is to adopt a neopositivistic or nominalistic attitude. That is not so, despite the fact that certain prominent logicians have represented and do represent this point of view. At all events a discussion of these views can only proceed on the basis of a general theory of knowledge; mathematical logic and, even more, research into the

foundations of mathematics can provide certain material for such a doctrine, but the proper discussion of the pros and cons of a given doctrine proceeds on other grounds. By and large we can also establish that all possible philosophical points of view are represented among mathematical logicians, naturally with the exception of those which contradict definite results of foundational research.

We come now to the objections to mathematical logic. A first objection, certainly to be taken seriously, is raised because various authors have set up calculi which have been called logical calculi yet for which no rational interpretation can be given. We have in mind here, in the first place, the so-called 'many-valued logics'. Whereas in ordinary logic the statements or the propositions are considered to be either true or false, one has here more than two and possibly infinitely many truth values. In so far as a version of the calculus of probability is under discussion there is nothing against this, but the expression 'many-valued logic' does not seem to me to be suitable. In so far as, on the other hand, one is dealing with calculi for which no interpretation whatsoever can be given I am in agreement with the critics in rejecting the view that these 'many-valued logics' have anything to do with logic as a philosophical discipline. It is not my intention to suggest that to concern ourselves with 'many-valued logics' is meaningless and useless. Such investigations may have their justification within the framework of a mathematics of calculi, but that is not a matter for philosophy. Nevertheless we must disallow the objection. After separating the philosophically irrelevant there remains a residue of true mathematical 'logic' which can and must also interest philosophers.

To avoid misunderstandings I wish to point out that neither the modal logics which we shall discuss later, nor intuitionist logic are many-valued logics in the sense used above.

A further objection dismisses mathematical logic with the remark that it just is not the same as the philosophical logic which forms the foundation of our thought. Mathematical logic, so it is said, may with its subtle definitions be a suitable instrument for investigating mathematical relationships, but for the rest we manage very well with the traditional Aristotelian theory of inference as a foundation for our entire scientific thought; even for mathematical inferences nothing else is necessary. This thesis is also presented in a recent book by v. Freytag-Löringhoff [6].[1] The merits of the traditional theory of inference are not denied here. Also there is no discontinuity with mathematical logic, for the traditional modes of inference can arise and be classified in it (cf. Hilbert-Ackermann [13], p. 44 ff). But we

[1] The numbers in square brackets refer to the bibliography at the end of the article.

must be clear that the Aristotelian modes of inference embrace only a part of our logical conclusions. Even if we disregard the fact that, prior to the construction of a mathematical logic, that part of logic which is concerned solely with statements as wholes and takes no account of their inner structure, the so-called statement logic, has not been set up or investigated at all, nevertheless the Aristotelian syllogism is insufficient where relations are concerned (cf. Hilbert-Ackermann [13], p. 49 ff). If a statement such as 'If a prime number different from two is odd and if eleven is different from two and a prime number, then eleven is odd' is correct on purely logical grounds, because the content of the concepts 'eleven', 'prime number', 'odd' and 'different from two' has no bearing on its correctness — the traditional syllogism of Barbara is expressed here — then not less so is the following statement: 'If in a town there is a chairman of all societies of the town, then every society of the town has a chairman.' Here also the content of the concepts 'chairman' and 'society of the town' is irrelevant to the correctness of the statement. But we are here concerned with the application of a proposition of the logic of relations which is not attainable with the Aristotelian modes of inference.

Of course one can make logical inference without knowledge of mathematical logic. Thus no logical formalism is used as a rule in the derivations of mathematics. But this certainly does not mean that Aristotelian logic is sufficient in mathematics; one needs relation logic just the same whether formalized or otherwise. The error in this respect can only have come about because nobody before Frege and Peano had undertaken a complete analysis of the modes of logical inference employed in mathematics. In older books on logic (cf. e.g. J. Fr. Fries [7]) one finds plenty of examples of the application of the traditional figures of inference in mathematics, and of complicated accumulations of these, but of no single important mathematical proposition is the proof dissected in complete detail. These modes of inference used in mathematics are none other than those employed in our scientific thought, except that complicated logical relationships rarely occur in non-mathematical fields of knowledge.

A formalization of logic has the same advantages over logic used purely intuitively as has formalization in mathematics. One can certainly derive mathematical theorems without using mathematical formulae, as was done by the ancient Greeks, but certain mathematical investigations, such as the theory of equations in algebra, are hardly thinkable without this aid. It is the same with the formalization of logic. This makes possible a precise formulation of logical inference free from all the ambiguities of common language, which

is indeed not exclusively a logical instrument. This precise formulation is also profitable for inference in mathematics, but here, where as a rule there are enough other controls on our thought, it is not so necessary as in other domains where emotional and other unconscious enthymemes threaten to falsify logical thought. Beyond this the formalization of logic makes it possible to undertake systematic investigations which are denied to purely intuitive thought; this is true, for example, of the problem of how to characterize the propositions which follow from given assumptions, or the problem of how one can in general determine whether a proposition is correct on purely logical grounds, the so-called decision problem. We shall have more to say later on the lack of appreciation shown by traditional logic of the difficulty which resides in the latter problem. At all events we can say in general that through the formalization of logic a swift development has occurred in a rather barren and apparently closed branch of knowledge. The Aristotelian syllogism does already represent an attempt, albeit imperfect, at formalization. One concluding remark: it should not be disputed that other problems concerning logic occur in philosophy with which mathematical logic is less occupied.

A further objection to mathematical logic is that it is incomplete, that it can only reproduce a part of our logical thought. This objection rests upon the following facts. In the year 1932 K. Gödel proved the theorem that the usual mathematico-logical systems are necessarily incomplete, i.e. theorems can be formulated using the formal means of expression of the system which are indeed correct from the intuitive point of view but for which no proof can be provided within the logical system. In each case the logical system can then in fact be extended so that the theorem in question is provable, but with the result that other correct theorems remain unprovable. We shall not give here an exact description of the logical systems comprehended by the Gödel theorem; suffice it to say that they are of a very general character. A. Church has shown in various writings, by related considerations, that within a formal system with sufficient means of expression the concept of 'consequence' cannot be completely defined. This necessary incompleteness of a formal system is now used against mathematical logic in that it is pointed out that non-formal, intuitive logical thought is complete.

In support of this argument, it is true that the entire wealth of our logico-mathematical thought cannot be reproduced within a single formal system. But to see a way out here by the rejection of formal systems would be wide of the mark. Naive intuitive logical thought, which has had its effect within mathematics in the formation of the

[illegible]ied naive set theory, has, if you like, completeness; here, however, one has also the logical and set-theoretical paradoxes (cf. in this connection, e.g. Grelling-Nelson [10]). We can also imitate naive logical thought in a formal system, but then contradictions appear as well. The Gödel theorem signifies more precisely: a formal system is either incomplete or not free from contradiction. Much has already been written on the paradoxes from the philosophical side — both illuminating and less illuminating. But by and large these considerations — I refer here only to the philosophers who are not informed about mathematical logic — miss the decisive point which is that these paradoxes arose unexpectedly, i.e. on the basis of modes of inference which from the naive intuitive point of view had to be considered correct. Thus a discussion of the paradoxes only has point if it is combined with a revision of the modes of inference employed, which have to be restricted so that no contradictions arise. This can be done in a precise way only through the setting up of a corresponding formal logical system; one would then similarly restrict intuitive inference so precisely in relation to the use of modes of inference that it would differ little from a formalized system, even though no logical symbols were used.

In regard to these necessary restrictions there is, up to now, no uniform, generally accepted recipe. This situation, incidentally, must not be understood to mean that arbitrary logical systems can now be set up whose meaning has first to be discovered, or whose meaning is supplied only by the prescribed usage. On the contrary, the meaning and usage of basic logical concepts such as 'and', 'or', 'not', and of universal and existential assertions, is fixed, and the fundamental part of logic which consists of the calculus of statements and the restricted predicate calculus, i.e. the theory of predicates and relations in a given domain of individuals, is also completely determined. An aberration of intuitionist logic in this respect is in principle not so significant, since this logic may still be interpreted in an intuitionist way. But we are concerned here with the further question: for which things in general is a defined predicate meaningful, i.e. for which things is it meaningful to say that the incidence of the predicate is true or false? From the naive point of view it is assumed that the incidence of a predicate is meaningful for every thing, or at least that with each predicate can be associated another which applies to the same things as does the first and which is meaningful for all things; then the paradoxes arise. In the so-called Theory of Types of Russell the idea is introduced that a predicate can never apply to itself, and in the axiomatic theory of sets likewise sets which can contain themselves as element are avoided. In both cases an

incompleteness is produced if certain propositions, about whose formulation there are no misgivings and which are clearly seen to be correct, are not provable in the system, or if it is nevertheless felt that the omission of all predicates which can be applied to themselves is a shortcoming. But it is a mistake to use this incompleteness against mathematical logic. One can say that this incompleteness applies even more to philosophy than to investigations into the foundations of mathematics, since here the system of the Russell Theory of Types, or at least that of axiomatic set theory, is entirely sufficient in practice for deriving mathematical theorems. In philosophy, on the other hand, considerations frequently occur in which concepts are applied to themselves without one being able to regard the considerations in question as incorrect, which suggests that the construction of a more extensive logic is necessary (cf. in this connection also F. B. Fitch [8]). We shall not report here upon the experiments that have already been undertaken in this connection. Final generality can hardly be achieved here. We have mathematical logic to thank for the discovery of this incompleteness, which is always present whether one employs a formalized logic or not.

We have to resign ourselves to the fact of the incompleteness in one respect or another of any precisely formulated logical system. The situation is strikingly reminiscent of the familiar antinomies set up by Kant. These antinomies depend in each case upon postulating, on theoretical grounds, a completeness in one direction or another which on the other hand can be proved not to hold good. Here we have a new antinomy, the antinomy of the completeness of logic.

It is clear that as a result of this a pure 'An-sich' interpretation of logic receives a distinct shock. By such an interpretation we understand one according to which the system of logical propositions exists independently of the thought we direct to it and needs only to be rediscovered. That is evidently not wholly the case; besides the thing given intrinsically an element of construction plays a part which can operate in various directions.

With this incompleteness of logic we have already indicated an important result of formalized investigation of foundations. In the enumeration of further results we shall deal only with those which touch on questions which have played or still play a part in a rational philosophy, particularly one connected with Kant. Problems of this kind are, for example, the investigation of the concepts 'necessity' and 'possibility'. C. I. Lewis [15], [16] introduced the treatment of these concepts in the calculus of logic. In the symbolic treatment of these concepts by him and his successors it appeared that they were by no means determined in absolute, unique fashion,

but that various interpretations presented themselves which could be defined with the help of symbolic logic (cf. e.g. O. Becker [2]).

In Kantian philosophy the concepts of analytic and synthetic judgment are particularly important. According to Kant a judgment is called analytic if its predicate is already contained in the concept of the subject; every other judgment is called synthetic. Connected with this classification is the Kantian view that the basic form of every judgment is the connection of one subject with one predicate; but in the more recent logic it has been shown that the union of several subjects by a relation occurs likewise as a basic logical form. A definition of analytic judgment which is also satisfactory from the standpoint of mathematical logic has been given by L. Nelson ([19], p. 792). He says: 'Analytic judgments will thus be those which arise from the mere form of the judgments. And the most general analytic judgments will be the fundamental postulates of logic.' However, for the practical application of this explanation, as Nelson himself mentioned, a more exact knowledge of the forms of judgments and logical modes of inference is necessary. In the sense of mathematical logic we would perhaps say: An analytic judgment is a universally valid judgment where the notion of universal validity is precisely determined within mathematical logic. I will not argue here against this division of judgments into analytic and synthetic, but only against a secondary sense which tends to be associated with 'analytic', and which is not present in Nelson's definition. The concept 'trivial', namely, has been associated with the concept 'analytic'. This may be because the examples cited as analytic judgments are mostly of the following kind: 'A stallion is male', where a stallion is then defined as a male horse, or: 'The triangle has three angles.' At all events one can often hear statements such as 'Logic is self-evident' and one meets with incredulous astonishment if one points out that it is an unsolved and, in its generality, not even a soluble problem to determine whether or not a given proposition is correct, i.e. analytic in the above sense, on purely logical grounds. In 1935 A. Church showed that there can be no general solution of this problem, which is known as the decision problem. The concept of a logical decision procedure, which in itself is somewhat vague, must of course for this purpose be defined in formal fashion. But it appeared that various precise renderings of this concept, undertaken from different starting points, coincided in essence, so that Church's result can be regarded as conclusive. Let it be noted further that this insolubility of the decision problem has been proved even for the so-called lower predicate calculus, a comparatively simple part of logic, concerning whose construction complete agreement prevails, as we have already noted.

Thus it is wrong to connect the concept 'trivial' with the concept 'analytic'. This triviality exists at most for the so-called propositional logic, i.e. for that part of logic in which no use is made of the analysis of propositions into subject and predicate or into subjects and relation, but in which propositions appear only as unanalysed wholes or, in so far as they are combined, are produced from other propositions by use of the propositional connectives 'and', 'or', 'not' and 'follows from'. Among the propositions considered here which are correct on purely logical grounds are found the following types: 'From the sentence A follows its double denial', 'If A is the case and from A B follows, then B is the case', 'If from the sentence A the sentence B follows and from the sentence B the sentence C, then also C follows from A'; in these the content of the sentences A, B, C does not matter. But even within the part of logic which corresponds roughly to the domain of application of the Aristotelian syllogism — we call it in mathematical logic the monadic predicate calculus — a solution of the decision problem is really difficult and was not achieved until the first decades of this century, while in general, as we have said, a solution of the decision problem cannot be contemplated.

That the determination whether or not a given sentence is correct on purely logical grounds is in general by no means trivial can be shown in the following way. Axiom systems for the theory of numbers can be set up, from which the statements of the theory of numbers can be logically derived; such an axiom system is, for example, the so-called Peano axiom system. We shall for the moment regard the axioms of these systems with Kant as synthetic statements. All derived sentences are then in general also synthetic, irrespective of the fact that they can be obtained from the axioms by logical inferences. Now a system of axioms does not give a complete definition of the concepts which occur in it, but only certain of their properties or their concatenation with other concepts. For example, it is specified of the fundamental number-theoretic successor relation that the successor of each number is uniquely determined, that there exists a number which is not a successor, and other things. But no more is required for the derivation of mathematical theorems either. If we have any mathematical sentence which follows from the axioms, then there corresponds to it a purely logical assertion whose correctness involves that of the mathematical sentence and conversely. If, for example, this mathematical sentence concerns some property of the successor relation, then this sentence is correct if and only if every relation which has the properties expressed in the axioms for the successor relation has also the properties attributed

in the mathematical sentence to the successor relation. But this is a purely logical assertion. Thus the correctness of every mathematical sentence, here, more precisely, every number-theoretic sentence, is equivalent to the correctness on purely logical grounds of another sentence. So a triviality of logic would carry with it a triviality of mathematics, and one is less inclined to accept the latter.

As concerns mathematics, it is significant that in some philosophical circles there still prevails a completely antiquated notion of it. Today one can no longer describe the substance of mathematics by saying only that it is the science of space on the one hand and of number on the other, as was true to a certain extent over a hundred years ago. The standard for our concept of mathematics is that which today is taught in colleges and published in scientific journals. Present-day mathematics investigates structures, abstract structures, one can also say logical structures. A structure is given by an axiom system in which appear certain concepts and relations which, however, are not completely defined, but of which only certain properties and possibly their interrelation with other concepts are specified. To give an example, let us consider the structure of a linear ordering. One is here concerned with a relation '$<$', besides '$=$', which satisfies the following axioms:

(1) For all x, y belonging to the domain of application of '$<$' we have either $x<y$ or $x = y$ or $y<x$.

(2) For all x and y it follows from $x<y$ that not $y<x$.

(3) For all x, y, z it follows from $x<y$ and $y<z$ that also $x<z$.

(4) For all x, $x = x$ holds.

(5) For all x and y it follows from $x = y$ that $y = x$.

(6) For all x, y, z it follows from $x = y$ and $y = z$ that also $x = z$.

(7) For all x, y, z it follows from $x<y$ and $x = z$ that also $z<y$.

(8) For all x, y, z it follows from $x<y$ and $y = z$ that also $x<z$.

Now what '$<$' or the things of the domain of application of '$<$' actually mean is not specified and is left entirely undetermined. For example, x, y, z could be events, '$x<y$' could be interpreted as 'the event x is earlier than the event y' and '$x = y$' as 'the event x is simultaneous with the event y'. But one could also give quite different interpretations, for example, that x, y, z denote natural numbers and that '$<$' denotes the relation 'less than'. In the investigation of the structure all these interpretations are disregarded, and it is also not even presupposed that x, y, z are specifically mathematical objects. The investigation of the structure now consists in determining which sentences can be derived from the axioms and which not, without laying down a particular interpretation of the axiom system.

But each derivability of this kind is equivalent to the purely logical correctness of a sentence, as we have already mentioned and now set forth more precisely. To give a simple example, the sentence 'For all x, not $x<x$' is derivable from the axiom system. Since no particular interpretation is employed in the derivation of the sentence, the derivability of the sentence can only mean the following:

For all relations R and S with the following properties:

(1) For all x, y either xRy or xSy or yRx holds;

(2) for all x and y it follows from xRy always that not yRx;

(3) for all x, y, z with xRy and yRz we also always have xRz;

(4) for all x we have xSx;

(5) for all x and y it follows from xSy that ySx;

(6) for all x, y, z it follows from xSy and ySz that also xSz;

(7) for all x, y, z it follows from xRy and xSz that also zRy;

(8) for all x, y, z it follows from xRy and ySz that also xRz,

it also holds that for all x not xRx. That is, however, a purely logical sentence. In the previous section we made the same statement about the Peano axiom system of the natural numbers.

If we now ask whether the axioms are analytic or synthetic sentences we are inclined to answer: 'synthetic', since the properties of '$<$' expressed in the axioms certainly do not apply to all relations. But we can only speak of analytic or synthetic when a definite interpretation is given. To the interpretation of x, y, z as events and '$<$' as 'earlier than' and '$=$' as 'simultaneous', mentioned previously, the concept synthetic applies. But let us now give an entirely different interpretation. By x, y, z we shall now understand the statements that one can form in any field of knowledge whatsoever. '$x<y$' shall mean 'The statement x is false and the statement y is true'. '$x = y$' shall mean 'The statements x and y are both true or both false', or briefly, 'both have the same truth value'. For this interpretation likewise all the axioms become correct sentences. For the following in fact hold:

(1) For all statements x and y, x is false and y true, or x true and y false, or both statements have the same truth value;

(2) If x is false and y true, then it is not the case conversely that y is false and x true.

(3) If x is false and y true and y false and z true, then also x is false and z true. (Here, by the way, the hypothesis of the statement cannot happen; independently of that the conclusion follows from it on purely formal grounds.)

(4) x has the same truth value as x.

(5) If x and y have the same truth value, then also y and x have the same truth value.

(6) If x and y have the same truth value and y and z have the same truth value, then also x and z have the same truth value.

(7) If x is false and y true and x and z have the same truth value, then also z is false and y true.

(8) If x is false and y true and y and z have the same truth value, then also x is false and z true.

However, these sentences have a purely logical character; they belong to the logic of statements. In the main, the only axioms of logic we use here are the one that a statement cannot at the same time be true and false, and the axiom of the excluded middle. We thus establish that it is possible for the axioms of a formal structure to be interpreted both by synthetic and by analytic sentences.

The conception of modern mathematics as the investigation of structure has emerged only in the last decades. A work is in course of preparation which will present the whole of mathematics from this point of view. It is the joint work of several French mathematicians who come together under the name Bourbaki. Examples of structures which are investigated in mathematics are group, ring, skew field, field, lattice, vector, module, ideal, topological space. What matters here is that the objects with which the structures deal are not necessarily specifically mathematical. A group may be concerned with natural or real numbers, geometric transformations, or quite different non-mathematical things. Consequently — and may this never happen in the average mathematical lecture — quite different structures can just as well be set up, such as the structure of biological relationships, in whose axioms relations and predicates like 'father of', 'mother of', 'ancestor of', 'male', 'female' play a part. Or one could set up the structure of marriage permission and prohibition according to the code of Civil Law by treating the laws in question in axiomatic form. At all events by turning to the investigation of structure, mathematics has entirely burst out of its former restricted framework. That a sentence can be derived from the axioms of a structure is, as we have seen, equivalent to the correctness of a definite logical sentence. Accordingly if we ask how this general mathematics differs from logic the answer runs: to a very slight degree, and mainly superficially. Mathematics in general employs no formalized logic and restricts itself to the investigation of certain structures which result from the generalization of those which already play a part in calculation with rational or real numbers or in ordinary geometry, and leaves aside the most general structure investigation of mathematical logic. But a borderline between this mathematics in the most general sense and mathematical logic cannot be drawn once and for all.

From the development of a general concept of mathematics given above it might appear that the specifically mathematical in the restricted sense had become superfluous. But this is not the case. Every occupation with structures, be they formulated with or without the help of mathematical logic, presupposes an intuitive operation with natural numbers or an equivalent operation with symbols. This is also the case when a structure of natural numbers is being dealt with. The obviousness connected with the concept of natural number is independent of all established structures. This obviousness lies in the fact that one conceives the notion of a unit by abstracting from the special qualities of a thing. Starting from one unit one can proceed indefinitely by appending more and more units, where the units can be represented materially in any way, for example, by strokes placed side by side, by points, by balls or the like. The sort of representation here is irrelevant, but one is not allowed to bring any special properties of the representative things into consideration. This conception enables us to comprehend as immediately obvious certain properties and relations for the natural numbers, for example, the less-than relation, and to realize other properties constructively or recursively, i.e. ascending from smaller to larger numbers. It is clear that this conception largely coincides with the corresponding Kantian views. A development of an intuitive theory of numbers of this sort has come from two sides, one from D. Hilbert and P. Bernays ([14], §2), the other within the school of Intuitionism founded by L. E. J. Brouwer (cf. A. Heyting [11]). There exist between the two schools, even as far as intuitive mathematics is concerned, yet finer distinctions of ideas which we shall not go into here. Besides the intuitive interpretation of mathematics there is something else common to both. In order to express complicated mathematical facts a logic is naturally employed. But this logic is only developed so far as is absolutely necessary for the connection of intuitive facts. Its (certainly rather superficial) distinction from the classical logic of the predicate calculus consists in the fact that the axiom of the excluded middle is no longer presupposed in full generality; for each assertion is conceived as a possible construction, and one cannot always say that the construction of an entity or the opposite construction succeeds. One must concede that this interpretation is consistent within itself. For the rest the distinction between the intuitionist and the classical theory of numbers is, at least in formal respects, not so significant. Above all, the two conceptions coincide for finite aggregates. Further, every classically correct statement is interpretable also as an intuitionistically correct statement. Admittedly this interpretation yields a different sense from that which

was originally intended. In the axiom of excluded middle, in the axiom that for every mathematical assertion A either A or not-A is true, the conviction is perhaps inherent that every mathematical problem is soluble, even if we do not know the solution at the moment. In the intuitionist interpretation the sentence would receive the meaning that A and not-A are not simultaneously the case.

The merit of Intuitionism as well as the Hilbert intuitive mathematics lies in having worked out a coherent minimal part of logic and mathematics, which restricts itself to that which is immediately intuitively understandable and to the modes of inference which lead from obvious to obvious, and for whose sentences in my view the character of sentences *a priori* in the Kantian sense may be upheld with good reason. Weighty differences between classical and intuitionist conception now arise as soon as the construction of real numbers is dealt with. One is here concerned less with different conceptions of mathematics as such than with the differences which arise during construction from the number-theoretic foundations in consequence of the various views on the extent of the logical concept formations, differences which we mentioned earlier when we spoke of a certain incompleteness of logic. If one adopts as basis the Russell Theory of Types in its simplified form, the theory of real numbers may be constructed in its classical form, as it is generally used in mathematics. But Intuitionism uses only the ramified logic of Types, incidentally first set up by Russell, so for it the classical theory of real numbers is no longer acceptable. Perhaps Intuitionism is right in its assertion that in the classical theory an arbitrary element of construction is added to the basic mathematical propositions. But the question arises whether one can do without such elements altogether. Finally, if one wishes to go into its foundations, the theory of real numbers can scarcely be constructed independently of geometric apperceptions. This is not unconnected with the question of whether or not there exists a geometric-intuitive conception of the continuum.

Of course the theory of real numbers in any form can also be constructed axiomatically as an abstract structure — and this happens as a general rule in mathematics. If one formalizes this structure with the help of mathematical logic, the intuitionist mathematics or the Hilbert intuitive mathematics (in the form of metamathematics) suffices for derivation from the axioms. But the setting up of a particular structure says nothing about its value for knowledge; because we can be satisfied with a generally acknowledged minimum of logic and elementary mathematics in the derivation of sentences from structures, we are able to work with arbitrary structures,

including those which we reject intuitively or about whose value for knowledge we have formed no opinion whatever. The various extensions of logic of which we spoke earlier, from the simple predicate calculus upwards, can in each case also be represented by a special structure which is formalized.

Throughout the preceding we took up the cause of an intuitive conception of the doctrine of the natural numbers. We saw further in the example of Intuitionism that it is possible to construct a mathematics independent of the customary logic, or, more simply, with a minimum of logic. It may seem that we are therefore forced to reject at once the logicism relating to the theory of numbers, which goes back to Frege and Russell. But it is not as simple as that. G. Frege [4], [5] and later, following his example, B. Russell (cf. [22]) undertook to define the notion of natural number purely logically and to derive from this definition all arithmetical propositions. We can make the following preliminary observation. For the natural numbers a definite structure can be specified, e.g. by the Peano axiom system mentioned earlier. Even if one takes the view that the assertions about natural numbers expressed in the axioms are synthetic propositions, it is still possible that there is at the same time a purely logical interpretation of the axiom system. In fact we had a case of this kind earlier as regards the structure of linear order. Consequently we can regard it as a matter of course that the structure of the Peano axiom system for the natural numbers admits both a mathematical and a logical interpretation, or, as we also say, that there exists for it a mathematical and a logical model.

But this does not adequately describe the situation produced by the Frege-Russell analysis of the number concept. If we call two predicates equinumerous when the things to which they apply can be set in a reversibly unique correspondence with one another, then by a number Frege understands a property of properties (a predicate of predicates), namely such a predicate as applies to all those and only those predicates which are equinumerous with one to which it applies. Within this general concept of number, which corresponds to the set theoretical concept of cardinal number, a narrower concept of inductive number may now be defined. By the number one we understand the predicate which applies to all and only those predicates which apply to only one thing. If this definition is called circular because the concept 'one' already appears in the definition of one, the answer is that all that matters is to obtain a purely logical definition of one. But this is the case. That a predicate applies to one and only one thing can be expressed with purely logical means, namely thus: 'There exists an x to which the predicate applies and

to all things different from x the predicate does not apply.' The concept of the successor of a number is likewise definable. By an inductive number is then understood a number to which applies every predicate which applies to one and to the successor of every number to which it applies.

If we now ask whether with the concept of inductive number we have attained that which we understand by natural number, we can hardly answer this question in the negative. For the refutation of the Frege-Russell theory points of view are usually cited which do not touch on the central core of this theory. These relate to difficulties which arise in the derivation of number-theoretic propositions from the definition mentioned, difficulties which are not concerned at all with this definition of number itself. Thus the Frege experiment failed in that it did not appreciate the paradoxical nature of a logic which imitates naive thought and led in consequence to contradictions. Russell was able to carry out his derivation by presupposing his Theory of Types. The chief objection to his derivation of the propositions of number theory is that he had to use a special axiom, the so-called axiom of infinity, which was not provable within his logic. But these things are not objections to the whole way of laying the foundation of the number concept; they only show that for deriving the mathematical propositions from the number concept one needs a particularly extensive system of logical modes of inference. W. V. Quine [21] in fact succeeded in deriving the mathematical propositions without an axiom of infinity. In this connection I may also refer to my work [1], in which likewise the purely logical derivation of the mathematical propositions is achieved. It is not essential that the concept 'set' is taken as fundamental in this work instead of the concept 'predicate' — 'set' results from 'predicate' by abstraction, in that one only distinguishes between predicates when they have different compass. At any rate one can no longer reject the Frege-Russell thesis on the logical nature of the number concept.

How does this result now stand in relation to our attitude presented earlier, which assumed an intuitive nature of the concept of natural number? Brouwer reduced his number intuition to the following concepts: the idea of a unity, which comes about by abstracting from the special qualities of any one object the possibility of distinguishing such a unity from another one and the possibility of arbitrarily repeating the second process. But in all these concepts there is nothing foreign to logic. One cannot construct logic at all without bringing in mathematical concepts as well. The number two is immanent in the concept of difference or non-identity, just as it is finally even in negation. For if a predicate applies to one object and not to another

these objects must necessarily be different from one another and represent the number two. If we take more predicates we can also characterize numbers such as three, four, etc., by their applying and not applying. Concepts such as 'order' and others which one usually regards as purely mathematical also play a part in logic. We should therefore not be surprised to hear that the development of the number concept within logic can be achieved in other ways besides that of Frege-Russell (cf. e.g. A. Church [3]). All in all we can say that in the theory of natural numbers we have a domain which is capable of an intuitive foundation with a minimum of logic in the sense of the statements on previous pages and also of attainment purely by logical definitions if we presuppose an extensive logic. This will appear surprising only to one who associates the concept of logical with that of trivial, which as we showed earlier is inadmissible.

Let us now turn to geometry, not, that is, to any abstract geometry, but to geometry in so far as it conveys to us a certain knowledge of the external world. It might appear superfluous to utter a single word about this. For the prevailing view is that among the theoretically possible abstract geometrical structures the one which is valid for the external world is wholly determined by physics. To a large extent that is indeed the case. But if one takes the view that this holds without exception one does not do justice to the fact that in our geometrical perception of space there also exist intuitive elements, intuitive impressions which force themselves upon us with compelling power. I am thinking here in the first place of certain topological impressions which are also upheld in the geometries which underlie the new physics, for example the proposition that a closed curve in a plane divides the region of the plane into two parts in such a way that we can only reach a point in one part from a point in the other part over a point of the curve. Moreover, the above-mentioned conception of geometry gives no explanation of the intuitive peculiarity of Euclidean space. On the other hand one also hears the view expressed that our entire representation of space, which is supposed to be that of Euclidean space, is given by our perception of space, as was Kant's opinion, but that — unfortunately — modern physics takes no account of this perception of space. That would seem to indicate a strange notion of the harmony of the different sides of our perceptive faculty.

It is not my intention to go into every detail of the involved relationships which exist here — a full account would also have to deal with the relation of our space representation to the visual and to the tactile space — if only for the reason that in my view there is much that has not been clarified. I will, however, make a few observations

which will allow the principal standpoint represented here to be understood. If one wishes to go more closely into the question of what is and what is not intuitively given in our space representation one cannot simply do it by taking some Axiomatics of Euclidean Geometry, for example that given by Hilbert [12], and then attempting by a kind of introspection to determine which of the axioms are intuitively evident and which are not. Apart from the possibility of a subjective error, there is usually in the formulation of the axioms a mixture of intuitive elements and those of an arbitrary construction. Let us consider, for instance, the concept of a straight line. Discussion in our sense about the axioms concerning the straight line has restricted itself essentially to the parallel axiom and to certain axioms of connection, such as whether two straight lines can always have at most one point of intersection. But contained in the concept of the straight line there is already that of the linear continuum, which in the Hilbert axiomatics is determined by the axioms of order and congruence for the straight line, by the Archimedean axiom and by the axiom of completeness. Do we now have an intuition of this continuum? Such an intuition is referred to by various authors. Brouwer, for instance, with his theory of real numbers strove, so he said, for an annexation to the geometrical theory of the continuum. It fits ill with this that his theory of real numbers has nothing to do with a Dedekind continuum, the best illustration of which is offered by his familiar theorem, false in ordinary mathematics, that every function defined throughout a closed interval is also continuous there. L. Nelson [18] says the intuitive concept of continuity of a curve would include that of differentiability. But since both continuity and differentiability contain the concept of limiting value and this bears on the continuum, one must ask oneself which intuitive concept of the continuum is to be taken as basic; for in ordinary mathematics the continuity of a function does not in fact entail its differentiability, as is shown by the familiar example of Weierstrass of a function which is everywhere continuous and nowhere differentiable. By a continuous curve in the intuitive sense one must understand one defined by an analytic function. But should the concept of an analytic function or that of a corresponding curve be intuitively comprehensible? Similar questions arise in respect of our representation of time; one can here indeed speak of a linear geometry. Prior to the introduction of choices of sequences in his theory of real numbers Brouwer considered the continuum as given through the intuition of time. It will be difficult for me to explain what is meant here. One could regard as immediately given that moments have a linear order in the sense of the postulates of order which we specified earlier, perhaps

also the metricizability in the sense both of the Archimedean axiom and of the Hilbert congruence axioms relating to the straight line. But how am I to perceive that the set of moments is everywhere dense, i.e. that between two moments there always lies another? That would presuppose a submicroscopically refined power of perception. A perception can, however, always relate only to something surveyable, finite. It is no contradiction to regard as possible an intuitive foundation of the doctrine of the natural numbers, for although certainly not every arbitrarily large number is intuitively comprehensible, this is true of the process of passing from one number to the next larger one, the constructive generative process of numbers. But one can, in a way, force intuition to the infinitesimal by postulating that every time segment is isomorphic to an arbitrary other one, i.e. that its moments can be mapped on to those of the other one in a reversibly unique manner so that the order and other relations are preserved. We could call this postulate that of complete intuition. Its introduction excludes the possibility that the moments form a discrete series like the natural numbers, however small the separation of the individual moments. But we must realize that this postulate of complete intuition is not itself of an intuitive nature, that it is not forced upon us by our intuition of space but involves an arbitrary element of construction which we have adopted to systematize our knowledge. Finally, in order actually to obtain the linear continuum, yet a further completeness postulate would be necessary, namely that which is called the Axiom of Completeness in Hilbert's *Grundlagen der Geometrie*.

Naturally the setting up of the time continuum can also be achieved from points of view which are entirely outside our intuition of time. If, say, we had reason to suppose that the points of space form a continuum, and further that every movement in space could only proceed continuously, then the continuum of time would thereby also be given. But we shall not go further into points of view of this sort.

The principle of complete intuition proves particularly efficacious in regard to our representation of space. We can express it here in such a way that an arbitrarily small piece of space would be isomorphic with an arbitrarily larger one of the same form, so that for instance an arbitrarily large cube would be indistinguishable from an arbitrarily small one as far as internal properties are concerned. Expressed in general terms this would mean that a part of space too large or too small for our comprehension can always be studied as a model of suitable size. If we now imagine a certain sequence of basic hypotheses concerning space to be given, in a similar manner as

for time perception, then it would follow from the postulate of complete intuition that space has a constant curvature, that the sum of the angles in a small plane triangle does not differ from that in a large one, and so on, i.e. we should have Euclidean geometry. This geometry is the only one which can be comprehended intuitively as a whole; in this lies its outstanding character. But we must realize that this complete perceptibility of Euclidean geometry is not given in itself but is enforced by a special postulate, by the postulate of complete intuition as we have called it. This postulate has not itself intuitive character, nor can it be described in any other way as *a priori*, but flows from the endeavour to round off our state of knowledge in a systematic way. It may be supplanted by other postulates, such as the postulate of combining our geometrical representations with our natural knowledge in a system as unified and simple as possible.

On the basis of our proceedings so far a certain conviction has impressed itself upon us, namely the conviction that certain facts of knowledge are given to us immediately or, as in arithmetic, are given to us directly through a constructive activity of our mind, but that other principles are needed as complementary to which an element of arbitrariness is attached and which no longer have this all-excluding character but can just as well be replaced by other principles. Prominent among these principles we have found completeness postulates. This notion is in sharp contrast with, on the one hand, a comprehensive Apriorism in the sense of Kant and, on the other hand, a thoroughgoing conventionalism whose motto is 'Truth by convention'. The objection which we could raise against the latter attitude, namely that its assumption would deprive our entire sense of truth of its proper foundation, would presumably make little impression; one would receive the reply that truth is indeed variable. But there is in my opinion a telling argument against that. Let us suppose any system of sentences to be axiomatically presented, together with the appropriate rules of inference, for which no generally acceptable interpretation can be given — say, a many-valued logic. Such a formulation of a system is indeed possible with the help of its symbolic formalization. The author now requires of us that we be able to verify and agree to the consequences which he draws from his system. That is only possible if he presupposes as self-evident to us such elementary logic and arithmetic as is used for operating with the symbols. To regard this as a mere result of primitive experimentation would scarcely be feasible, for its obligatory character could then hardly be explained.

If we have adopted the conception presented above, there arises

the important question of how we can determine which part of our knowledge can be regarded as generally obligatory and for which part that is not the case. We have here to deal with an uncertainty factor, particularly because in our mathematical statements the intuitively given is nearly always intermingled with the result of free constructions. As we see no way of solving this problem, it may be useful to survey in the history of philosophy the treatment of questions of this kind in logico-mathematical or related fields. It is clear in the first place that this question can only have been treated in a philosophy which is especially concerned with the *a priori*, with generally obligatory items of knowledge. This occurs in the philosophy of Kant and in that of J. Fr. Fries, who extended and developed the Kantian philosophy. The problem of determining which of our items of knowledge is generally obligatory and which is not is called, at least in the Kant-Fries philosophy, the problem of philosophical deduction (cf. L. Nelson [17]). We seek in vain, however, in that philosophy for accounts of logico-mathematical 'deduction' – this concept has naturally nothing to do with what one calls deduction in logic. The apriority of logic, arithmetic and Euclidean geometry seemed to be so exalted above all doubt that a deduction in this domain was considered almost superfluous and was never seriously sought, though on occasion it was mentioned as a problem. But in another domain, that of the philosophical founding of physics, serious efforts have been made to find a philosophical deduction of certain metaphysical principles, e.g. the axiom of reciprocity, the axiom of causality and the axiom of conservation of matter. As we are looking for methods, and as, on the other hand, logic plays a part in the trains of thought to be described, the short digression here into this non-mathematical field may be justified.

For the deduction of the above-mentioned metaphysical principles Kant – and, following his example, Fries – used the so-called transcendental guide. Kant considered that human understanding could be conscious of no other basic metaphysical concepts than those which he conceived through the logical forms of judgments. He concluded further that each basic form of logical judgment must also correspond to a definite metaphysical basic judgment. This view was to a certain extent confirmed because there existed a far-reaching parallelism between the logical basic forms as he understood them and the metaphysical basic judgments as they were then generally acknowledged following the current Newtonian physics. Thus for him, through the transcendental guide, i.e. through the complete enumeration of the logical basic judgments, there was given at the same time a complete enumeration of the metaphysical basic judgments

and these were thereby proven as such. For this consideration, naturally, all propositions which contain a combination of concepts are only rendered by means of the judgment forms. But even if the Kantian enumeration of the logical basic forms of judgments were complete — which in the light of modern logic is doubtful — it is still a logically arbitrary assumption that each logical basic form must correspond to exactly one metaphysical basic judgment. This is not affected by any parallelisms which may exist between logic and metaphysics.

But Kant had another principle for discovering metaphysical judgments; he said they were principles of the possibility of experience, i.e. principles which when applied to empirical sense-data make possible scientific judgments of experience. Principles of this kind can be obtained by the analysis of our statements about our experience (cf. L. Nelson [20]). However, such an abstraction from experience shows only the rational character of these principles, but leaves open the question of whether we are not here concerned with arbitrary constructions. Thus in order to establish the generally obligatory character of principles of this kind one has yet to show that they cannot be dispensed with, that they cannot be replaced by other principles.

Now it appears to me that this second method of Kant's can also be fruitful for our field as well. Let us illustrate what we mean by this for geometry. A short time ago we presented the view that certain basic statements of geometry are generally obligatory, while for completeness other principles are to be appended in which free constructions must occur, whether one is to be guided in the setting up of these principles by points of view within geometry or by consideration of the whole system of our natural knowledge. This view must not be misunderstood as meaning that certain geometrical principles must be forced dictatorially upon physics. This case can hardly arise, for if it is clear that one can manage in physics without certain geometrical propositions it is thereby proved that they are not generally obligatory. We have here an exclusion principle. Such exclusion principles are provided not only by present-day physics but by any possible physics in which one considers not only physical systems of the past but also new developments, merely possible, in fancy, and without reference to their experimental confirmation. Geometrical propositions which underlie every possible physics should be regarded as generally obligatory. Naturally not all possible physical systems are at our disposal. In practice, for progress in this field we have in the main to fall back on the fact that individual inspired architects of science construct systems by means of which certain

hitherto generally accepted geometrical axioms are then shown to be dispensable.

Thus the problem of laying down the generally obligatory principles in geometry is not one of which one can say with absolute certainty that it is finally resolved at a certain point in time. At least that is the case so long as one has not found a special philosophical deductive process, and I would simply not be able to say anything about such a process. I would nevertheless not conclude that the number of those geometrical propositions which are possibly generally obligatory would converge to zero with the advance of science; but I am conscious that the argument against such a view is only a psychological one, namely the fact of the impressiveness and intuitive nature of certain geometric representations. At all events it is no contradiction to assume that certain geometric representations are generally obligatory, and at the same time to assume that we cannot determine with absolute certainty the scope of this condition. In the logico-arithmetical field it is my opinion that the apriority of the intuitive elementary doctrine of numbers together with the appertaining logic can in practice be regarded as secured.

PAPERS AND BOOKS REFERRED TO IN THE TEXT

[1] W. Ackermann, 'Zur Axiomatik der Mengenlehre', *Math. Annalen*, 131 (1956), pp. 336-45.

[2] O. Becker, *Untersuchungen über den Modalkalkül*, Meisenheim 1952.

[3] A. Church, 'Mathematical Logic', Mimeographed, 1935.

[4] G. Frege, *Grundlagen der Arithmetik*, 1884.

[5] G. Frege, *Grundgesetze der Arithmetik*, 1893-1903.

[6] B. von Freytag-Löringhoff, *Logik — ihr System und ihr Verhältnis zur Logistik*, Stuttgart 1955.

[7] J. Fr. Fries, *System der Logik*, Heidelberg 1837, Reprint Leipzig 1914.

[8] F. B. Fitch, 'Self-reference in philosophy', *Mind*, n.s., vol. LV (1946), pp. 64-73.

[9] K. Gödel, 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme', *Monatshefte für Mathematik und Physik*, 37 (1931), pp. 173-83.

[10] K. Grelling und L. Nelson, 'Bemerkungen zu den Paradoxien von Russell und Burali-Forti', *Abhandlungen der Friesschen Schule*, Neue Folge, II Bd., Heft 3, 1908.

[11] A. Heyting, *Intuitionism. An introduction*, Amsterdam 1956.

[12] D. Hilbert, *Grundlagen der Geometrie*, 8 Aufl., Stuttgart 1956.

[13] D. Hilbert und W. Ackermann, *Grundzüge der theoretischen Logik*, 3 Aufl., Berlin-Göttingen-Heidelberg 1949.

[14] D. Hilbert und P. Bernays, *Grundlagen der Mathematik*, I Bd., Berlin 1934.

[15] C. I. Lewis, *A Survey of Symbolic Logic*, Berkeley 1918.

[16] C. I. Lewis and C. H. Langford, *Symbolic Logic*, New York 1932.

[17] L. Nelson, 'Die kritische Methode und das Verhältnis der Psychologie zur Philosophie', *Abhandlungen der Friesschen Schule*, Neue Folge, I Bd., 1 Heft, Göttingen 1904.

[18] L. Nelson, *Des fondements de la géometrie*, Göttingen 1918.

[19] L. Nelson, 'Über das sogennante Erkenntnisproblem', *Abhandlungen der Friesschen Schule*, Neue Folge, II Bd., 4 Heft, Göttingen 1908.

[20] L. Nelson, 'Ist metaphysikfreie Naturwissenschaft möglich?' Ibid., II Bd., 3 Heft, Göttingen 1908.

[21] W. V. Quine, *Mathematical Logic*, 2nd ed., Cambridge (Mass.) 1951.

[22] B. Russell, *Introduction to Mathematical Philosophy*, 3rd ed., Cambridge 1924.

THE AIM OF SCIENCE*

By Karl R. Popper

To speak of 'the aim' of scientific activity may perhaps sound a little naive; for clearly, different scientists have different aims, and science itself (whatever that may mean) has no aims. I admit all this. And yet it seems that when we speak of science, we do feel, more or less clearly, that there is something characteristic of scientific activity; and since scientific activity looks pretty much like a rational activity, and since a rational activity must have some aim, the attempt to describe the aim of science may not be entirely hopeless.

I suggest that it is the aim of science to find *satisfactory explanations*, of whatever strikes us as being in need of an explanation. By an *explanation* (or a causal explanation) is meant a set of statements one of which describes the state of affairs to be explained (the *explicandum*) while the others, the explaining statements, form the 'explanation' in the narrower sense of the word (the *explicans* of the *explicandum*).

We may take it, as a rule, that the *explicandum* is more or less well known to be true, or assumed to be so known. For there is little point in asking for an explanation of a state of affairs which may turn out to be entirely imaginary. (Flying saucers may represent such a case: the explanation needed may not be of flying saucers, but of the reports of flying saucers; but should flying saucers exist, then no further explanation of the *reports* would be required.) The *explicans*, on the other hand, which is the object of our search, will not be known as a rule: it will have to be discovered. Thus, scientific explanation, whenever it is a discovery, will be *the explanation of the known by the unknown*.[1]

The *explicans*, in order to be satisfactory (satisfactoriness may be a matter of degree), must satisfy a number of conditions. First it must logically entail the *explicandum*. Secondly, the *explicans* ought to be true, although it will not, in general, be known to be true; in any case, it must not be known to be false even after the most critical examination. If it is not known to be true (as will be usually the case) there must be *independent* evidence in its favour; or in other words, it must be *independently* testable; and it will be the more

* This is a section from the author's *Postscript: After Twenty Years* to his *The Logic of Scientific Discovery*, about to be published. A few references to other parts of the book have been omitted here.

[1] See the last paragraph of my text (before the final quotation) of my 'Note on Berkeley as a Precursor of Mach', *Brit. Journ. Philos. Sc.* 4, 1953, p. 35.

satisfactory the more independent and the more severe the tests were to which it has stood up.

I thus have still to elucidate my use of the expression 'independent', with its opposites, '*ad hoc*', and (in extreme cases) 'circular'.

Let *a* be an *explicandum*, known to be true. Since *a* trivially follows from *a* itself, we could always offer *a* as an explanation of itself. But this would be highly unsatisfactory, even though we would know in this case that the *explicans* is true, and that the *explicandum* follows from it. *Thus we must exclude circular explanations of this kind.*

Yet the kind of circularity I have here in mind is a matter of degree. Consider the following dialogue: 'Why is the sea so rough today?' — 'Because Neptune is very angry today' — 'By what evidence can you support your statement that Neptune is very angry today?' — 'Oh, don't you *see* how *very* rough the sea is? And is it not always rough when Neptune is angry?' This explanation is found unsatisfactory because (just as in the case of the fully circular explanation) the only evidence for the *explicans* is the *explicandum* itself. The feeling that this kind of almost circular or *ad hoc* explanation is highly unsatisfactory, and the demand to avoid explanations of this kind are, I believe, among the main motives for the development of science: they are the first fruits of the critical or rational approach.

In order that the *explicans* should not be *ad hoc*, it must be rich in content: it must have a variety of testable consequences, and among them, more especially, testable consequences which are different from the *explicandum*. It is these different testable consequences which I have in mind when I speak of *independent* tests, or of *independent* evidence.

Although these remarks may perhaps help to elucidate a little the intuitive idea of an independently testable *explicans*, they are still quite insufficient to characterize a satisfactory and independently testable explanation. For if *a* is again our *explicandum* — such as 'the sea is rough today' — then we can always offer a highly unsatisfactory *explicans* which is completely *ad hoc* even though it has independently testable consequences. We can still choose these consequences as we like. We may choose, say, 'These plums are juicy' and 'All ravens are black'. Let *b* be their conjunction. Then we can take as *explicans* simply the conjunction of *a* and *b*: it will satisfy all our requirements so far stated.

Only if we require that explanations make use of universal statements or laws of nature (supplemented by initial conditions) can we make some progress towards realizing the idea of independent, or non-*ad hoc*, explanations. For universal laws of nature *may* be statements with a rich content, so that *they may be independently tested*

everywhere, and at all times. Thus if they are used as explanations, they *may* not be *ad hoc* because they *may* allow us to interpret the *explicandum* as an instance of a reproducible effect. All this is only true, however, if we confine ourselves to universal laws which are testable, that is to say, falsifiable.

The question, 'What kind of explanation may be *satisfactory*?' thus leads to the reply: an explanation in terms of testable and falsifiable universal laws and initial conditions. And an explanation of this kind will be the more satisfactory the more highly testable these laws are (and the initial conditions too), and the better they have been tested.

In this way, the conjecture that it is the aim of science to find satisfactory explanations leads to the further idea of improving the degree of satisfactoriness of the explanations by improving their degree of testability, i.e. by proceeding to better and better testable ones; which means proceeding to theories of richer and richer content: of higher degree of universality, and of higher degree of precision.[2] This, no doubt, is a result fully in keeping with the actual practice of the theoretical sciences.

We may arrive at fundamentally the same result also in another way. If it is the aim of science to explain, then it will also be its aim to explain what so far has been accepted as an *explicans*, such as a law of nature. Thus the task of science constantly renews itself. We may go on for ever, proceeding to explanations of a higher and higher level of universality — unless, indeed, we should arrive at an *ultimate explanation*; that is to say, at an explanation which is neither capable of any further explanation, nor in need of it.

Are there ultimate explanations? The doctrine which I have called 'essentialism' upholds the view that science must seek ultimate explanations in terms of essences:[3] if we can explain the behaviour of a thing in terms of its essence — of its essential properties — then no further question can be raised, and needs to be raised (except perhaps the theological question of the Creator of the essences). Thus Descartes believed that he had explained physics in terms of the *essence of a physical body* which, he taught, was extension; and some Newtonians, following Roger Cotes, believed that the *essence of matter* was its inertia and its power to attract other matter, and that Newton's theory may be derived from, and thus ultimately

[2] For the theory of *testability*, *content* and *simplicity*, and of degrees of *universality* and *precision*, see sections 31 to 46 of my *Logic of Scientific Discovery*, where the close connection between these ideas is explained.

[3] I have discussed (and criticized) essentialism more fully in my paper 'Three Views Concerning Human Knowledge', where I also refer to my earlier discussions (in the last footnote to Section ii); see *Contemporary British Philosophy*, iii, edited by H. D. Lewis, 1956, note 2 on p. 365.

explained by, these essential properties of all matter. Newton himself was of a different opinion. It was a hypothesis concerning the ultimate or essentialist causal explanation of gravity itself which he had in mind when he wrote in the *scholium generale* at the end of the *Principia*: 'So far I have explained the phenomena . . . by the force of gravity, but I have not yet ascertained *the cause of gravity itself* . . . and I do not arbitrarily [or *ad hoc*] invent hypotheses.'[4]

I do not believe in the essentialist doctrine of ultimate explanation. In the past, critics of this doctrine have been, as a rule, instrumentalists: they interpreted scientific theories as *nothing but* instruments for prediction, without any explanatory power. I do not agree with them either. And, indeed, there is a third possibility, a 'third view', as I have called it. It has been well described as a 'modified essentialism'—with emphasis upon the word 'modified'.[5]

This 'third view' which I uphold modifies essentialism in a radical manner. First of all, I reject the idea of an ultimate explanation: I maintain that every explanation may be further explained, by a theory of higher universality; and there can be no explanation which is not in need of a further explanation, for none can be a self-explanatory description of an essence (such as an essentialist definition of body, as suggested by Descartes). Secondly, I reject all *what-is questions*: questions asking what a thing is, what its essence is, or its nature. For we must give up the view, characteristic of essentialism, that there is an essence, an inherent nature of principle in every thing (such as the spirit of wine in wine) which causes it, or necessitates it to be what it is, and thus to act as it does. This animistic view explains nothing; but it has led essentialists (like Newton) to shun relational properties, such as gravity, and to believe, on grounds felt to be *a priori* valid, that a satisfactory explanation must be in terms of inherent properties (as opposed to relational properties). The third and last modification of essentialism is this. We must give up the view, closely connected with animism (and characteristic of Aristotle as opposed to Plato), that it is the essential properties inherent *in each individual or singular thing* which may be appealed to as explaining this thing's behaviour. For this view completely fails to throw any light whatever on the problem why different

[4] See also Newton's letters to Richard Bentley of January 17th and especially February 25th, 1693 ('1692-3'). I have quoted from this letter in Section iii of my paper 'Three Views Concerning Human Knowledge', where the problem is discussed a little more fully.

[5] The term 'modified essentialism' was used as a description of my own 'third view' by a reviewer of my paper 'Three Views Concerning Human Knowledge' in *The Times Literary Supplement*, **55**, 1956, p. 527. In order to avoid misunderstandings, I wish to say here that my acceptance of this term should not be construed as a concession to the doctrine of an 'ultimate reality', and even less as a concession to the doctrine of essentialist definitions. I fully adhere to the criticism of this doctrine which I have given in my *Open Society*, Chapter 11, Section ii (especially note 42), and in other places.

individual things should behave in like manner. If it is said, 'because their essences are like', the question arises *why there should not be as many different essences as there are different things*. (Plato tried to solve precisely this problem by saying that like individual things are offsprings, and thus copies, of the same original 'form', which thus is something 'outside' and 'prior' and 'superior' to the various individual things; and indeed, we have as yet no better theory of likeness. Even today, we appeal to their common origin if we wish to explain the likeness of two men, or of a bird and a fish, or of two beds, or two motor cars, or two languages, or two legal procedures; that is to say, we explain similarity in the main genetically; and if we make a metaphysical system out of this, it is liable to become a historicist philosophy. Plato's solution was rejected by Aristotle; but since Aristotle's version of essentialism does not contain even a hint of a solution, it seems that he never grasped the problem.[6])

By choosing explanations in terms of universal laws of nature, we offer a solution to precisely this last (Platonic) problem. For we conceive all individual things, and all singular facts, to be subject to these laws. The laws (which in their turn *are* in need of further explanation) thus explain regularities or similarities of individual things or singular facts or events. And these laws are not inherent in the singular things. (Nor are they Platonic ideas outside the world.) Laws of nature are conceived, rather, as (conjectural) descriptions of the structural properties of nature — of our world itself.

Here, then, is the similarity between my own view (the 'third view') and essentialism: although I do not think that we can ever describe, by our universal laws, an *ultimate* essence of the world, I do not doubt that we may seek to probe deeper and deeper into the world or, as we might say, into properties of the world that are more and more essential.

Every time we proceed to explain some conjectural law or theory by a new conjectural theory of a higher degree of universality, we are discovering more about the world, trying to penetrate deeper into its secrets. And every time we succeed in falsifying a theory of this kind, we make a new important discovery. For these falsifications are most important. They teach us the unexpected; and they reassure us that, although our theories are made by ourselves, although they are our own inventions, they are none the less genuine assertions about the world; for they can *clash* with something we never made.

Our 'modified essentialism' is, I believe, helpful when we raise the

[6] As to Plato's theory of forms or ideas, it is 'one of its most important functions to explain the similarity of sensible things . . .'; cp. my *Open Society*, Chapter 3, Section v; see also notes 19 and 20, and text. The failure of Aristotle's theory to perform this function is mentioned there (in the third edition, 1957) at the end of note 54 to Chapter 11.

question of the logical form of laws of nature. It suggests that our laws, or our theories, must be *universal*, that is to say, must make assertions about the world — about all spatio-temporal regions of the world. It suggests, moreover, that our theories make assertions about structural or relational properties of the world; and that the properties described by an explanatory theory must be, in some sense or other, deeper than those to be explained. I believe that this word 'deeper' defies the attempt at an exhaustive logical analysis; it is, however, a guide to our intuitions. (This is so in mathematics: all its theorems are logically equivalent, in the presence of the axioms, and yet there is a great difference in 'depth' which is hardly capable of logical analysis.) The 'depth' of a scientific theory seems to be most closely related to its simplicity and so to the wealth of its content. (It is otherwise with the depth of a mathematical theorem whose content is always nil.) Two ingredients seem to be required: a rich content, and a certain intuitive coherence or compactness (or 'organicity') of the state of affairs described. It is this latter intuitive ingredient which is so difficult to analyse, and which the essentialists were trying to describe when they spoke of essences, in contradistinction to a mere accumulation of accidental properties. I do not think that we can do much more than refer here to an intuitive idea. And I do not think that we need to do much more. For in the case of any particular theory proposed, it is the wealth of its content, and thus its degree of testability, which decides its interest, and the results of actual tests which decide its fate. From the point of view of method, we may look upon its depth, its coherence and its beauty as a mere guide or stimulus to our intuition and to our imagination.

Nevertheless, there does seem to be something like a *sufficient condition* for depth, or for degrees of depth, which can be logically analysed. I shall try to explain this with the help of an example from the history of science.

As everybody knows, Newton's dynamics achieved a unification of Galileo's terrestrial and Kepler's celestial physics. It is often said that Newton's dynamics can be induced from Galileo's and Kepler's laws, and it has even been asserted that it can be strictly deduced from them.[7] But this is not so; from a logical point of view,

[7] What can be deduced from Kepler's laws (see Max Born, *Natural Philosophy of Cause and Chance*, 1949, pp. 129-33) is that, for all planets, the acceleration towards the sun equals at any moment k/r^2, where r is the distance at that moment between the planet and the sun, and k a constant, the same for all planets. Yet this very result formally contradicts Newton's theory (except on the assumption that the masses of the planets are all equal or, if unequal, then infinitely small as compared with that of the sun). This fact follows from what is here said, in the text following footnote 9, about Kepler's third law. But in addition, it should be remembered that neither Kepler's nor Galileo's theories contain

Newton's theory, strictly speaking, contradicts both Galileo's and Kepler's (although these latter theories can of course be obtained as approximations, once we have Newton's theory to work with). It is therefore impossible that Newton's theory can be obtained from these theories by a process of either deduction or of induction. For neither a deductive nor an inductive inference can ever lead from consistent premises to a conclusion which formally contradicts these premises.

I shall now briefly indicate the contradictions. Galileo asserts that a thrown stone or a projectile moves in a parabola, except in the case of a free vertical fall when it moves, with constant acceleration, in a straight line. (We neglect air-resistance throughout this discussion.) From the point of view of Newton's theory, these assertions are both false, for two distinct reasons. The first is false because the path of a long-range projectile, such as an inter-continental missile, will be not even approximately parabolic: it is elliptic. It becomes, approximately, a parabola only if the total distance of the flight of the projectile is negligible, compared with the radius of the earth. This point was made by Newton himself, in the *Principia* as well as in its popularization, *The System of the World*, where he illustrates it with the help of the figure reproduced on the opposite page.[8]

Newton's figure illustrates his statement that, by increasing the velocity of the projectile, and thus the distance of its flight, it will, 'at last, exceeding the limits of the earth, . . . pass into space without touching it'.

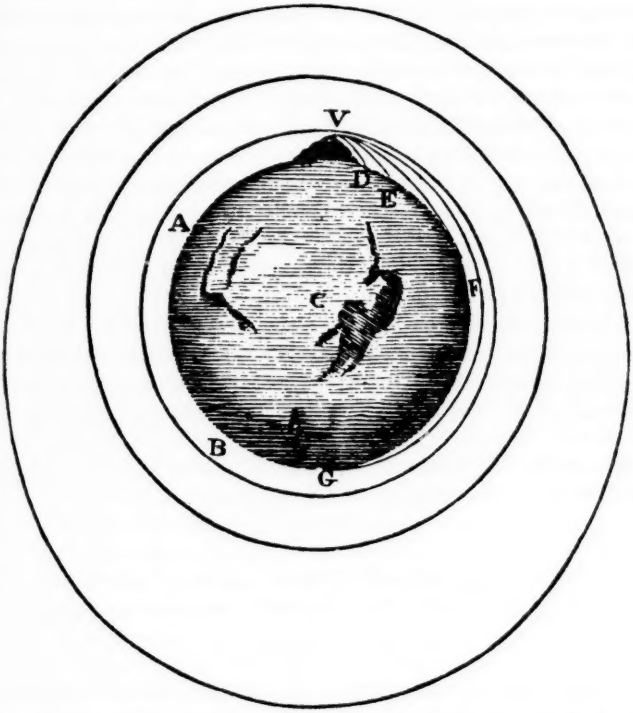
Thus all projectiles on earth move along ellipses with finite eccentricities. Of course, for sufficiently short throws, Galileo's parabola will be an excellent approximation; but the parabolic track is not strictly deducible from Newton's theory unless we add to the latter a factually *false* initial condition (and one which, incidentally, is unrealizable in Newton's theory) to the effect that the radius of the earth is infinite. If we admit only initial conditions which are not known to be false, then we always get an ellipse with a finite eccentricity,

[8] See Newton's *Principia, Scholium* at the end of Section ii of Book i; p. 55 of the 1934 edition (Motte's translation revised by Cajori). The figure, from *The System of the World*, and the quotation following it, will be found on p. 551 of this edition.

Newton's concept of *force* which is traditionally introduced in these deductions without further ado; as if this ('occult') concept could be read off the facts, instead of being the result of a new interpretation of the facts (that is, of the 'phenomena' described by Kepler's and Galileo's laws) in the light of a completely new theory. Only *after* the concept of force (and even the proportionality of gravitational and inertial mass) has been introduced is it at all possible to link the above formula for the acceleration with Newton's inverse square law of attraction (by the assumption that the planets' masses are negligible).

which contradicts Galileo's law according to which we should obtain a parabola.

A precisely analogous logical situation arises in connection with the second part of Galileo's law which asserts the existence of an



acceleration *constant*. From the point of view of Newton's theory, the acceleration of free falling bodies is never constant: it always increases during the fall, owing to the fact that the body approaches nearer and nearer to the centre of attraction. This effect is very considerable if the body falls from a great height, although of course negligible if the height is negligible, as compared with the radius of the earth. In this case, we can obtain Galileo's theory from Newton's if we again introduce the *false* assumption that the radius of the earth is infinite (or the height of the fall zero).

Both the contradictions which I have pointed out are far from negligible for long-distance missiles. To these we may apply Newton's theory (with corrections for air-resistance, of course) but not

Galileo's: the latter leads simply to false results, as we may show with the help of Newton's theory.

With respect to Kepler's laws, the situation is similar. It is obvious that Kepler's laws are only approximately valid — that is, strictly invalid — in Newton's theory if we consider the mutual attraction between the planets.[9] But there are more fundamental contradictions between the two theories than this somewhat obvious one. For even if, as a concession to our opponents, we neglect the mutual attraction between the planets, Kepler's third law, considered from the point of view of Newton's dynamics, cannot be more than an approximation which is applicable to a very special case: to planets whose masses are equal or, if unequal, negligible as compared with the sun. Since it does not even approximately hold for two planets one of which is very light while the other is very heavy, it is clear that Kepler's third law contradicts Newton's theory in precisely the same sense as Galileo's.

This can be easily shown as follows. Newton's theory yields for a two-body system — a binary star system — a law which astronomers often call 'Kepler's law' since it is closely related to Kepler's third law. This so-called 'Kepler's law' says that if m_0 is the mass of one of the two bodies — the sun, say — and if m_1 is the mass of the other body — a planet, say — then, choosing appropriate units of measurement, we can derive from Newton's theory

$$(1) \qquad a^3/T^2 = m_0+m_1$$

where a is the mean distance between the two bodies, and T the time of a full revolution. Now Kepler's own third law asserts that

$$(2) \qquad a^3/T^2 = \textit{constant},$$

that is to say, the same constant for *all* planets of the solar system. It is clear that we obtain this law from (1) only under the assumption that m_1 is the same for all planets; or, if this is factually *false* (as is indeed the case, since Jupiter is some million times larger than the smallest planets), that the masses of the planets are *all zero as compared with that of the sun*, so that we may put $m_1 = 0$, *for all planets*. This is quite a good approximation from the point of view of Newton's theory; but at the same time, putting $m_1 = 0$ is not only strictly

[9] See, for example, P. Duhem, *The Aim and Structure of Physical Theory*, 1905; English translation by P. P. Wiener, 1945, Part ii, Chapter vi, Section 4. Duhem says more explicitly what is implicit in Newton's own statement (*Principia*, Book i, proposition lxv, theorem xxv) for Newton makes it quite clear that in case that more than two bodies interact, Kepler's first two laws will be at best only approximately valid, and even this in very special cases only, of which he analyses two in some detail. Formula (1), below, follows immediately from Book i, proposition lix, in view of Book i, proposition xv. (See also Book iii, proposition xv.) Thus my own analysis, like Duhem's, is also implicit in Newton's.

speaking false, but unrealizable from the point of view of Newton's theory. (A body with zero mass would no longer obey Newton's laws of motion.) Thus even if we forget all about the mutual attraction between the planets, Kepler's third law (2) contradicts Newton's theory which yields (1).

It is important to note that from Galileo's or Kepler's theories we do not obtain even the slightest hint how these theories would have to be adjusted — what false premises would have to be adopted, or what conditions stipulated — in order to proceed from these theories to another and more generally valid one such as Newton's. *Only after we are in possession of Newton's theory can we find out in which sense the older theories were approximations to it.* We may express this fact briefly by saying that although from the point of view of Newton's theory, Galileo's and Kepler's are excellent approximations to certain special Newtonian results, Newton's theory cannot be said, from the point of view of the other two theories, to be an approximation to their results. All this shows that logic, whether deductive or inductive, cannot possibly make the step from these theories to Newton's dynamics.[10] It is only ingenuity which can make this step. Once it has been made, Galileo's and Kepler's results may be interpreted as corroborating the new theory.

Here, however, I am not so much interested in the impossibility of induction as in *the problem of depth.* And regarding this problem, we can indeed learn something from our example. Newton's theory unifies Galileo's and Kepler's. But far from being a mere conjunction of these two theories — which play the part of *explicanda* for Newton's — *it corrects them while explaining them.* The original explanatory task was the deduction of the earlier results. It is solved, not by deducing them, but by deducing something better in their place: new results which, under the special conditions of the older results, come numerically very close to these older results, and at the same time correct them. Thus the empirical success of the old theory may be said to corroborate the new theory; and in addition, the corrections may be tested in their turn — and perhaps refuted, or else corroborated. What is brought out strongly, by the logical situation which I have sketched, is the fact that the new theory cannot possibly be *ad hoc* or circular. Far from repeating its *explicandum*, the new theory contradicts it, and corrects it. In this way, even the evidence of the *explicandum* itself becomes independent evidence for the new theory. (Incidentally, this analysis allows us to *explain the value of metrical theories*, and of measurement; and it thus helps us to

[10] The concepts of force (cp. footnote 7, above) and of action at a distance introduce further difficulties.

avoid the mistake of accepting measurement and precision as ultimate and irreducible values.)

I suggest that if in the empirical sciences a new theory of a higher level of universality successfully explains some older theories *by correcting them*, then this is a certain sign that the new theory has penetrated deeper than the old ones. The demand that a new theory should contain the old one approximately, for appropriate values of the parameters of the new theory, may be called (following Bohr) the '*principle of correspondence*'.

Fulfilment of this demand is a sufficient condition of depth, as I said before. That it is not a necessary condition may be seen from the fact that Maxwell's electromagnetic wave theory did not correct, in this sense, Fresnel's wave theory of light. It means an increase in depth, no doubt, but in a different sense: 'The old question of the direction of the vibrations of polarized light became pointless. The difficulties concerning the boundary conditions for the boundaries between two media were solved by the very foundations of the theory. No *ad hoc* hypotheses were needed any longer for eliminating longitudinal light waves. Light pressure, so important in the theory of radiation, and only lately determined experimentally, could be derived as one of the consequences of the theory.' This brilliant sketch, due to Einstein,[11] of some of the major achievements of Maxwell's theory as compared with Fresnel's, may be taken as an indication that there are other sufficient conditions of depth not fully covered by my analysis.

The task of science, which, I have suggested, is to find satisfactory explanations, can hardly be understood if we are not realists. For a satisfactory explanation is one which is not *ad hoc*; and this idea — the idea of *independent evidence* — can hardly be understood without the idea of discovery, of progressing to deeper layers of explanation; without the idea, therefore, that there is something for us to discover; and that there is something to be critically discussed.

And yet, it seems to me that within methodology we neither have to presuppose metaphysical realism nor can we derive any help from it, except of an intuitive kind. For once we have been told that the aim of science is to explain, and that the most satisfactory explanation will be the one that is most severely testable and most

[11] A. Einstein, *Physikalische Zeitschrift*, **10**, 1909, p. 817 f. The abandonment of the theory of a material ether (implicit in Maxwell's failure to construct a satisfactory material model of it) may be said to give depth, in the sense analysed above, to Maxwell's theory as compared with Fresnel's; and this is, it seems to me, implicit in the quotation from Einstein's paper. Thus Maxwell's theory in Einstein's formulation is perhaps not really an example of *another* sense of 'depth'. But in Maxwell's own original form it is, I think.

severely tested, we have been told all we need, as methodologists. That the aim is realizable we cannot assert, neither with nor without the help of metaphysical realism which can give us only some intuitive encouragement, some hope, but no assurance of any kind. And although a rational treatment of methodology may be said to depend upon an assumed, or conjectured, aim of science, it certainly does not depend upon the metaphysical and most likely false assumption that the true structural theory of the world (if any) is attainable for men, or expressible in human language.

If the picture of the world which modern science draws is anywhere near to the truth — that is to say, if we have anything like 'scientific knowledge' — then the conditions pertaining almost everywhere in the universe make 'scientific knowledge', that is to say the discovery of structural laws of the kind we are seeking, almost impossible. For almost all regions of the universe are filled by chaotic radiation, and almost all the rest by matter in a likewise chaotic state. In spite of this, science has been miraculously successful in proceeding towards what I propose that we consider as its aim. I do not think that we can explain this strange fact without proving too much. But it can offer us encouragement to pursue this aim — even though we get no further encouragement to believe that we can actually realize it, neither from metaphysical realism nor from any other sources.

RELIGION, EXPERIENCE AND METAPHYSICS

By Julius Kraft

I. THE TWO BASIC PROBLEMS OF RELIGIOUS PHILOSOPHY

'Every mortal soul believes either in God or in an idol.' This powerful dictum occurs in Scheler's book *On the Eternal in Man*, which appeared in 1921. The generation of the First World War and of the period that followed was deeply impressed by Scheler's call. But the Second World War, and many of the events preceding it, have so numbed and hardened the human mind that it reacts only to the exaltation of religious sensation and is hardly receptive to the serenity of genuine contemplation.

Yet no thinking person can help but notice to what an extraordinary extent history has confirmed Scheler's dictum. For what else but idolatry was the racial doctrine of National Socialism and what else is the class doctrine of Communism? The generation of the Second World War has had to suffer the consequences of this political make-believe, idolatry and devilry. While lamentations, however pious, will not help us avoid shocks like these, the very memory of them could spur us on to a deeper understanding of the causes of the folly that leads men to commit such crimes against humanity. It is closely, although not irrevocably, bound up with human nature itself, and by no means restricted to man's public life. His private life is permeated by such folly which presents a great obstacle to knowledge of oneself and of others. But few people have the courage, and fewer still the lucidity and power of expression, to give as frank an account of their own and others' self-delusions as did Abélard in his reflections on Héloïse or Shakespeare in his 62nd Sonnet.

Scheler, like Abélard, belongs to the Augustinian, that is to say, the Platonic tradition. Thus the religio-philosophic appeal in *On the Eternal in Man* is based on the self-confidence of a metaphysical theory of truth formulated as intuitive certainty. But this claim to certainty cannot be maintained, since it goes beyond the bounds of human nature which is much more correctly characterized by the words in 2 Corinthians v. 7: '. . . we walk by faith and not by sight'. Even so, Scheler's *On the Eternal in Man* is a landmark in the twentieth-century reawakening of religious philosophy. The significance of Scheler's contribution cannot be dismissed by giving it such superficial historistical labels as, for instance, 'Catholicizing Platonism'. In religious philosophy, Scheler has taken up Husserl's

call 'to the nature of things themselves'. Instead of historicizing about this call, we should follow it whole-heartedly; only lack of judgment or blindness can lead us to deny the continued need for fundamental religio-philosophic clarification, when even the most elementary principles still await careful and precise analysis. Attempts to solve the fundamental problems of religious philosophy without such 'religio-philosophic prolegomena' merely demonstrate a lack of understanding of the difficulties involved, and no edifying assurances can explain away this fact.

These difficulties are grouped around two main problems: (1) the problem of the content of religious philosophy, i.e. what does religious philosophy assert? and (2) the problem of the validity of its assertions, i.e. what is the truth upon which its assertions are based?

In the twentieth century, both problems have been attacked with arguments which, taken to their logical conclusion, would amount to the annihilation of rational religious philosophy. No one wishing to continue its thousand-year-old tradition can ignore these attacks without exposing himself to the accusation that he is wearing philosophical blinkers.

The following remarks are not meant to tackle the two main problems of religious philosophy systematically, but

(1) To show that rational religious philosophy need not fear to face its critics, and

(2) To formulate certain conditions which, although not sufficient for the solutions of the two main problems, are, nevertheless, essential for them.

The first group of problems leads to a discussion of the religio-philosophic consequences of modern empiricism, of neo-orthodoxy and of existentialism. The second group leads to a more specific differentiation between religious life, religious knowledge, and religious philosophy, together with a clearer definition of religious philosophy itself, which involves a formulation, essential to the whole of philosophy, of the postulate of autonomous religious philosophy and an examination of its practical significance.

II. SCEPTICAL FORMS OF MODERN RELIGIO-PHILOSOPHIC EMPIRICISM

White spoke of the warfare between theology and science. This war has ended in an indisputable victory for science. No theological rearguard actions or occasional sorties can alter this fact.

The age-long struggle between theology and science having been decided, a new conflict has broken out over the relationship between religion and experience. In the name of empirical science — viz.

psychology and sociology — the whole of religion (and not only its dogmatics) is dismissed as an illusion. Thus the suggestive title of Freud's psychological version of modern religious scepticism is *The Future of an Illusion*, and Max Weber's sociological version speaks of the demystification of social reality and upholds a religious relativism which is allegedly connected with it.

These empirically important and original approaches show great intellectual honesty and courage, thus setting an example to the philosopher who tries to assess their significance. More than anything, they throw new light on the psycho-pathological element in man's religious life, and on the close interrelationship between religious, economic and political institutions. Both approaches are evidence of a substantial deepening of cultural and historical realism, but neither they nor similar viewpoints (e.g. those of Pareto and Frazer, or of comparative anthropology) have in fact the religious and philosophical consequences usually claimed for them. These consequences follow only from the fact that their authors, quite wrongly, regard certain philosophical assumptions as self-evident. Thus Freud, for instance, thought that his own sound rationalism ('There is no appeal beyond reason', equivalent to Kant's 'Quod supra nos, nihil ad nos') implied the empiricist conclusion that religious ideas, not being deposits of experience, must be illusions due to the pathological fantasies of a humanity in the grip of a general compulsion neurosis.

What is above all striking in this train of thought is the fact that it is clearly aware of the non-empirical character of religious ideas. Its conclusion would be valid if indeed there were no difference between the non-empirical and the counter-empirical.

It is widely believed that psychoanalysis has provided an exact basis for religious scepticism. But this belief is untenable. Freud's *Future of an Illusion* is an amazingly stimulating improvisation which, in its own way, fully recognizes the lasting importance of man's religious need which must be satisfied for the sake of his psychological balance. Yet it fails to integrate these psychological needs into the system of his Libido theory. For Freud's attempt to explain religion as a result of the father fixation — an attempt which is productive if carried out in a critical manner — is basically circular since a religious element is implied in the father fixation itself.

Max Weber is Freud's counterpart in the sociological sphere. Above all, his criticism of the materialistic interpretation of history was constructive because it demonstrated the vital importance of ethical and religious beliefs to the functioning of social systems. This analysis, conducted on a world-wide historical scale, shows

both the ramifications of, and the changes in, such socially effective ideas, thus implying that such ideas are only relative. In truth, the reverse is the case, for the very changeability of these ideas assumes the existence of constant and specific psychological dispositions. Admittedly, the sociology of religion is religiously and ethically neutral: to use Max Weber's term, 'valuefree'. However, this applies equally in the negative sense, i.e. it has no sceptical implications. Sceptical consequences arise merely from failure to distinguish adequately between religion and magic. Only if all religion were magic would the removal of the magical element mean the uprooting of religion itself.

Thus Max Weber's religio-philosophical scepticism, far from being the result of his sociology of religion, stems from a certain lack of conceptual clarity, bound up with a deep-rooted historistical approach. This approach fails to distinguish adequately and consistently between the unchanging basis of human nature and its variable overt forms of expression.

Max Weber's spiritualistic historicism and Freud's naturalistic positivism are both forms of modern empiricism and, hence, of empiricism in general. It is of great philosophical importance to realize clearly that modern religio-philosophic scepticism — no matter whether it is based on psychology or sociology — is in actual fact the old ethical scepticism of Hume joined historically, though not logically, to the recent emergence of new hypotheses and new branches of empirical science. There is no reason, therefore, for declaring war on psychology and sociology in the name of religion, nor for fighting religion in the name of psychology and sociology. But religious philosophy, in defining its fundamental attitude towards empiricism, must seriously consider the negative religio-philosophic consequences of modern empiricism. It must reformulate its questions with a view to meeting the objections of the latter and must avoid riding rough-shod over experience.

III. DOGMATIC FORMS OF MODERN RELIGIO-PHILOSOPHIC EMPIRICISM

This warning is by no means directed exclusively to positive religion, which has always shown a tendency to meddle in the sphere of experience. As Justice Holmes so rightly said of William James's religious philosophy, modern empiricism in its turn has helped 'to open the door wide for superstition'.

This striking and unexpected result of the empiricist approach to philosophy is logically unavoidable. In its examination of religio-philosophic concepts, empiricism comes to two conclusions, both

having the empiricist axiom itself as their first premise. If this premise, i.e. the proposition that 'all knowledge is empirical knowledge', is joined to the negative proposition that religious knowledge is not empirical knowledge, we have religio-philosophic scepticism, but if it is joined to the proposition that religious knowledge does exist, then there follows the empirical character of religious knowledge and thus a form of religio-philosophic mysticism.

In his original and influential book *The Varieties of Religious Experience* (1902), William James has systematized empiricist religio-philosophic mysticism. However, his theory of religious experience (just like Rudolf Otto's theory of the numinous) is based on a fundamental error: James and Otto both mistake psychological and historical approaches to the eternal for religious experience. Both of them make religious matters the subject of psychological and historical experience, without in any way establishing genuine religious experiences.

James and Otto described a series of significant facts about religious life, but these were complex, and therefore requiring further analysis. In particular, they have demonstrated that the religious horizon stretches far beyond the limits of institutional religion. However, they have not succeeded, nor indeed could they have succeeded, in proving any genuine intuition of the eternal – the self-interpretations of so-called religious seers not being anything like sufficient evidence for its existence.

The visions of such seers are known to differ a great deal. Philosophically and historically, the most important of these differences are those between seers who have visions of the world beyond and those whose visions extend only to this world. While James and Otto deal mainly with the first type, Dewey in *A Common Faith* claims to be a secular visionary. In fact all he has done is to reconstruct Rousseau's doctrine of 'La volonté générale'.

It is obvious, however, that these philosophers felt uneasy about their 'knowledge' of these religious experiences. Thus they introduced a most typical and far-reaching supplementary theory, according to which it is not knowledge but a life-promoting 'will to believe' (James's own term) which establishes the validity of religion's claim to truth. With this the transition has been made from empirical religio-philosophic mysticism to religious pragmatism (whose psychotherapeutic form has been put forward by Jung). While this pragmatism (or instrumentalism), with its emphasis on the will, is a form of decisionism, i.e. of authoritarianism, its emphasis on life-promotion makes it a form of utilitarianism (transferred from ethics to the theory of knowledge).

Now, the sceptical and mystical forms of religio-philosophic empiricism have retained the primacy of knowledge, but from their pragmatic approach there follows the primacy of the will, which means in fact the elimination of the concept of truth. For this reason alone, pragmatism is completely useless as a foundation of religious philosophy. In neo-pragmatic efforts to replace the concept of knowledge by that of 'performance', pragmatism inevitably loses itself in pedantic or playful constructions of religious performance. An example of what can more rationally and unpretentiously be accomplished in this direction is Huizinga's *Homo Ludens*.

The linguistic philosophers have fared no better than the pragmatic and neo-pragmatic philosophers. In their endeavour to make a virtue of the non-sense consequence of their dogmatic criterion of meaning, they believe to have found in the alleged non-sense of religious assertions the true profundity of these assertions. The old *credo quia absurdum* appears here to be resurrected by linguistic means, and this historical reference may suffice to characterize this pseudo-modern illogicality.

Compared with neo-scepticism, the dogmatic assertions of modern religio-philosophic empiricism prove to be desperate and fragile attempts at a reconciliation between empiricism and religion. Such a reconciliation is basically impossible and whoever attempts it is ignoring the commandment 'not to make unto thee any graven image'.

Philosophically speaking, this condemnation of idolatry means that no sound religious philosophy can be built on an empiricist foundation. Rational religious philosophy agrees with both forms of empiricism in that it assumes that religious knowledge exists and that it is non-empirical. However, it avoids the sceptical, mystical and pragmatic implications of these assumptions by refusing to restrict knowledge to experience and by reassigning to reason its place in religion.

IV. FORMS OF MODERN RELIGIO-PHILOSOPHIC IRRATIONALISM

The first of these assumptions is made by modern religious irrationalism which exists in a theological as well as a philosophic form, particularly (but by no means exclusively) in Barth's dialectical theology and in the religious philosophy of existentialism, fluctuating as it does between atheism and the most varied kinds of ecclesiastical beliefs.

What revelation is for dialectical theology, 'subjectivity', in all its forms, is for the existentialist philosopher. Yet revelation, as well as

subjectivity, represents a criterion of truth which contradicts the empiricist axiom. Beyond that (and this is more serious), they also reject the proposition that the counter-experiential and the incomprehensible are not admissible. It is in this, their paradoxical character, that dialectical theology and existentialism believe they possess their most valuable asset. According to a presumptive aphorism of Kierkegaard, a thinker without paradox is a 'mediocre fellow'. Every rational religious philosopher, to whom freedom from contradiction is important, must needs accept the fact that he will be labelled 'mediocre' by the Kierkegaardians. But surely the 'mediocrity' of Leibniz, Kant and Brentano are more valuable than any amount of neo-orthodox or existentialist 'profundity'.

This preference, however, must not prevent the rational philosopher from trying to fathom the recurrence of religious irrationalism in the twentieth century. Without such an understanding there can be no constructive conquest of religious irrationalism (or of irrationalism in general).

Now, we must acknowledge the fact that there is an element of honesty and consistency also in modern religious irrationalism. When Kierkegaard and Barth storm against the historicizing, psychologizing and philosophizing pseudo-interpretations of revelation and oppose them with their *sola fide*, *sola gratia*, then, from their own points of view, they are both consistent and honest. When Sartre, the empiricist and ethical anti-supernaturalist, rejects religion as well as morals, he too is consistent and honest. Rational religious philosophy must treat religious irrationalism with the same lack of equivocation, theoretically and practically. Both cannot have an equal right to a place in the sun. When the Gospel says that the last shall be the first, these words await their fulfilment as far as philosophy is concerned.

A mere glance at religio-philosophic discussion at the beginning of the nineteenth century convinces one of the guiding and anticipatory function of philosophy, at least as regards the appearance of religious irrationalism in the twentieth century.

At that time Hegel in his *Lectures on the Philosophy of Religion* was constructing a history of religion based on the so-called auto-motion of ideas. Christianity in particular is looked upon by Hegel as the 'absolute religion', the 'completed religion', the 'objectivized spirit of being itself'. In his *Philosophy of Mythology*, Schelling, arguing against this incredible verbiage, points out that all being contains a large share of unreasonableness as well, and that it is impossible to construct the existence and content of a positive religion dialectically. Although in speculative fantasy Schelling's theogony is hardly an

improvement on Hegel's ideophany, Schelling has already rejected Hegel's speculations about positive religion. Kierkegaard and his theological followers are even more radical. Not only do they follow Schelling in rejecting Hegel's logicistic philosophy, but they reject all philosophic knowledge in general. Neo-orthodoxy pleads exclusively for revelation, and existentialism for 'existence' which cannot be reduced to knowledge.

Both pleas, however, rest on a common elementary fallacy. The rejection of Hegelian philosophy does not inevitably lead to the rejection of all philosophical knowledge, unless it is assumed that what philosophical knowledge there is has been discovered by Hegel. Only on this assumption would the realization that Hegel did not possess the philosopher's stone actually mean the death sentence for philosophical knowledge. Kierkegaard's (and his followers') rejection of rational philosophy is, therefore, a consequence of their hypothetical Hegelianism. While they no longer believe that the world spirit speaks through Hegel, they do believe that, this not being the case, philosophical knowledge is denied to man. Kierkegaard's religion of paradox, far from being an elementary assertion unconnected with philosophy, is actually a form of pseudo-religious positivism based on elementary philosophical errors — a positivism which eschews thought precisely where it is most needed.

This is also the case with the rejection of rational religious philosophy on the part of dialectical theology. As Kierkegaard rightly opposes Hegel's metaphysics, so Barth opposes Harnack's historicist construction of religion. Barth, while taking up Lessing's assertion that truth concerning the eternal can never be founded on history, has discarded Lessing's and Harnack's respect for an autonomous history of religion which is independent of religion. Barth's approach is strictly ecclesiastical and authoritarian, and thus an unhistorical treatment of Protestant tradition.

Furthermore, this dogmatic approach is comparatively undifferentiated. This is evident from the fact that instead of learning from the demythologizing efforts of Bultmann, it attacks them with fanatical zeal. In view of these fundamental weaknesses, rational religious philosophy has no reason to feel intimidated by the self-assurance and the innumerable volumes which dialectical theology possesses.

This does not hold for Barth alone, but equally well for Tillich, Jaspers and Sartre.

However, in contrast to Barth's solid doctrinaire approach, these existentialists are infected with theological doubts. Tillich tries to save himself from this predicament by arguing theology, philosophy,

history, sociology, psychology and (in the sense of religious Socialism) politics, all in one.

His consistent 'correlation' of dogmatics and existentialism greatly helps this versatility, since the latter sanctions the illogicality of the former philosophically. All those who regard this theologico-philosophical double-talk as an absurdity are simply rebuked for their lack of dialectical understanding (which 'betrays' their character as bourgeois ideologists).

Jaspers, although not a theologian, is at one with Tillich in raising ambiguity to a principle. His philosophy has a counter-thesis for every religio-philosophic thesis, while it is never clear to what extent he is putting forward a neutral historical report or a thesis on dialectical tensions. Although less rigid denominationally than Tillich, Jaspers, following Schelling, opposes Bultmann's demythologization of theology in the name of his mystical myth-theory. In so doing, he provides philosophic reinforcements for Barth's neo-orthodoxy (while Tillich recognizes Bultmann's merits and appropriates them).

Because of Jaspers's and Tillich's sympathetic attitude to social, political and cultural progress and because of their intellectual ubiquity, their theology and philosophy are very tempting for some groups with an appropriate bias. But this attraction rests on a considerably weaker foundation than that of the uncompromising *credo quia absurdum*.

While Tertullian limits his dogma of absurdity to theology, Sartre has raised it to the level of universal principle by asserting that all things are incidental and 'therefore' absurd. This 'therefore' rests on an elementary logical trick, on the unfounded hyper-logicistical assertion: 'What is not logically necessary is absurd.' True, the (wrongly asserted) accidental nature of all things then leads to the principle of universal absurdity, and the very principle of a casual universe precludes the possibility of necessity and *a fortiori* of an absolutely necessary being.

Furthermore, Sartre's atheism rests on another ground which is important to our understanding of the religio-philosophic situation of the twentieth century. He looks upon his religious (and moral) nihilism as a simple consequence of his critique of miraculous, religio-moral theologizing authoritarianism. In fact, this consequence implies the additional assumption that religion and morals are identical with that authoritarianism. As Kierkegaard's rejection of philosophy rests on a hypothetical Hegelianism, so Sartre's rejection of religion and morals rests on a hypothetical theologism. His negativistic radicalism is radical in appearance only,

and is no serious argument whatsoever against rational religious philosophy.

Thus reviewing twentieth-century religious irrationalism, we discover that it is a house built on sand. Although it keeps mouthing the great word *Truth* it displays a sovereign contempt for reason and for the biblical exhortation that man's communication be yea, yea or nay, nay. Instead it has raised paradox, i.e. the yea-nay, to the level of a sacred shrine, thus belying its outward show of inner certainty.[1]

The religio-philosophical situation created by modern empiricism and irrationalism is a religious vacuum. Only a new understanding of 'religion within the boundaries of mere reason' can fill this vacuum.

V. RELIGIOUS LIFE, RELIGIOUS KNOWLEDGE AND THE PHILOSOPHY OF RELIGION

As a first step towards such an understanding, some basic distinctions must be made within the sphere of religion itself, distinctions which as a rule are blurred.

Religion cannot be identified with a particular church or a particular catechism, nor is it equivalent to the ideas about the ultimate held by poets and philosophers. Religion is a characteristically broad term with at least three meanings: religious life, religious knowledge, and the systematization and rational foundation of this knowledge, viz. the philosophy of religion. It is true that the religious life cannot be reduced to purely intellectual functions, but is the result of the interaction of all man's inherent spiritual tendencies. But it always presupposes a religious conviction which, in whatever form it may appear, must have a philosophical orientation. Thus there is an essential connection between religious life and the philosophy of religion, which does not at all arise from an attitude of religious intellectualism. Nor does the philosophy of religion claim to have a monopoly of religious truth. On the contrary, it readily grants that much of this truth is contained in the religious avowals of artists and in the feelings of simple unspoiled human beings. However, religious philosophy alone can lead to logical clarity and verification of the highest truths to which men can aspire. It alone can safeguard these truths against the distortions to which they are always prone. These distortions affect the whole of religious life as well as religious knowledge, and 'religion within the bounds

[1] This is discussed in greater detail in my booklet, *Knowledge and Belief*, Leiden, 1937.

of mere reason' opposes them not only by stipulating how religion is to be lived, but also by preserving the purity of its ideas.

Religious life and religious knowledge have always been classified under such sacrosanct but quite inadequate headings as 'experience' and 'faith'. True, there are certain psychological processes which correspond to these aestheticized (or dogmatized) headings, but these processes ought to be explained without such vague circumlocutions, i.e. consistently and empirically. If this is done, it becomes clear that although there are no religious experiences — not only must we make no false images, but we can never even form a real image — religion, in both thought and practice, does in fact begin with experience without however springing from it. It also becomes obvious that there are typical stages in the religious reaction to experience — e.g. momentary surrender to an overwhelming impression, particular rituals, and finally the transcendence of the sensory element in free thought. Yet this freedom is not only a source of emancipation, but also one of profound spiritual and social enslavement. Primitive men think that they are magically transformed if they disguise themselves and act accordingly. On a more subtle level, equally magical rituals make use of imaginative ideas and words. The dogmas, rituals and external disguises of Hitler's short-lived 'Thousand-year Reich' illustrate this ageless technique in most recent history. In contrast to this self-flattering 'magic transformation' of man's inner and outer life, it is the task of the philosophy of religion constantly to make clear man's apartness from the eternal and the perfect. Thus it cannot help taking up the struggle against human presumption which hides under the cloak of sham piety. This has been its inexorable position ever since Anaxagoras, Erasmus, Spinoza and Kant. To steer clear of this struggle, and to follow Hegel into a metaphysical construction of the positive, thus elevating philosophy, and particularly one's own, to divine worship, means turning magician and abandoning the role of philosopher.

The opposite and equally untenable extreme is to deny all reality to religious concepts, conceding to them no more than a consolatory value. Even this concession is somewhat dubious, for how can religion bring solace to people who are convinced of its illusory character?

Nevertheless, religious solace does exist even for the most knowledgeable of us, and this is so because religion, while starting from experience, actually points beyond experience. Only the philosophy of religion can explain this transcendence with the necessary clarity and lucidity.

VI. RELIGIOUS PHILOSOPHY AS TRANSCENDENTAL METAPHYSICS

This can be done by following different paths. One of them is the familiar distinction between higher and lower religions, which may be based, for instance, on an assessment of the artistic level of a particular cult, of the ethical level of the moral precepts in a particular theology, or of the stage reached in the development of religious ideas. In this respect there is a tremendous difference between, for instance, the concept of the divine in the Book of Genesis and in the Book of Job, the former being considerably more anthropomorphic than the latter. Here it is assumed that the further a religion is from anthropomorphism, the more highly developed is its concept of the divine, the majesty of the Creator being irreconcilable with the limitations of human nature, or of any other natural phenomena such as the sun, or even with nature itself. For this reason, and because of the autonomy of nature, philosophy cannot admit either natural or supernatural attributes of God in any but a symbolic sense. This applies to all religious ideas: their subject-matter cannot be characterized within either a natural or a supernatural sphere. For their subject-matter is characterized by the eternal aspect of the universe.

Greatly improved methods of thought are needed for an exact introduction, into philosophy, of the idea of the eternal order of things. This introduction cannot be effected by a general notion of metaphysics or by a nominal definition of ontology. Only by contrasting natural (i.e. temporal) and absolute, timeless, and thus eternal order, can we do justice to this task. It is insufficient simply to characterize the philosophy of religion as metaphysics. The philosophy of religion is metaphysics in an exalted sense, and cannot be described either as natural philosophy or as ethics. It does not restrict itself to any set of temporal phenomena, but is concerned with the eternal order of the universe. The philosophy of religion is transcendental metaphysics, in other words the metaphysics of an order which transcends nature, in contradistinction to such immanent metaphysics as natural philosophy and ethics which are concerned with natural order. No philosophy can avoid defining its attitude towards transcendental metaphysics; in fact, the struggles raging about it are the vital nerves of all serious philosophic efforts, no matter whether the protagonists deem transcendental metaphysical truths attainable or not.

Undoubtedly the problems of transcendental metaphysics are the most difficult and profound problems encountered in the whole of philosophy. Unfortunately, instead of being taken for philosophy's

ultimate concern, they have always been held to be its first task, with the inevitable result that they have remained unresolved. But there are exceptions to this rule, i.e. religio-philosophical attempts that do justice to the greatness of their subject-matter. Thus, in his critique of the traditional proofs of the existence of God, Kant, while irrefutably showing that there are inherent philosophical limitations to such proofs, distinguished between things as phenomena and as entities with intrinsic properties. By formulating this critical or transcendental idealism, he developed an idea which is the key to the solution of the religio-philosophical problems, and thus to the harmonious co-ordination of religio-philosophic principles with immanent metaphysics and with the total experience of nature. In fact, a religious philosophy cannot be considered adequate if it cannot be reconciled with the most radical naturalistic approach to the reality of experience. Critical idealism leads to such a reconciliation by dethroning naturalism as a universal principle and restricting it to the role of a principle of interpreting experience. This is no arbitrary compromise, but a consequence of the theory of the antinomies which shows that the usual attitude to reality, according to which the world of space and time is the world as it intrinsically exists, leads to contradictions. Faced with this fundamental logical objection, 'naive realism' must renounce its *naïveté* and make way for an 'empirical realism' that accepts the phenomenal character of the world of space and time.

The distinction between the phenomenal and noumenal aspects of the universe and its validation by the theory of the antinomies raises to conceptual clarity what we naturally feel about the limits of human reason and the symbolics of higher religion. Despite this fundamental accord between the boldest philosophic abstraction and our innermost feelings, the resulting philosophic outlook on the noumenal aspect of the universe is unsatisfactory. What real gains there are, are obscured by disturbing traditional notions in the 'Critique of Pure Reason' itself, by subsequent metaphysical re-interpretations and by re-presentations devoid of all philosophic insight. Moreover, further clarification and re-application are essential, and this is a very important and difficult task. Friedrich Schiller states this with sublime aesthetic clarity when, in his poem *The Three Words of Faith*, he declares: 'High above space and all time soars the loftiest thought.' Despite Kant's theory of the antinomies, the logical clarity of transcendental metaphysics has hardly anything equivalent to offer. It has still to prove itself through the exact systematization of the ideas of religion (free from all Church dogmas). It can only pass this test if it ceaselessly faces up to the

most challenging logical, metaphysical and mathematical abstractions bordering on its own field (e.g. Cantor's significant distinction between the transfinite and the absolute). The philosophy of religion is not a sermon; it must be approached in the spirit which Maimon rightly demanded for philosophy as a whole — the spirit of mathematical detachment, of objectivity instead of subjectivity. Objectivity allows of no contradictions, and thus rational philosophy of religion rejects all attempts to reintroduce transcendental ideas into psychological, historical, sociological and perhaps even physical experiences. They are impermissible not only from an empirical point of view, but also from the point of view of transcendental metaphysics itself. All such attempts fall far short of the present (if far too often ignored) position of critical religious philosophy which at least realizes clearly that its domain is not of this world.

VII. THE POSTULATE OF AN AUTONOMOUS PHILOSOPHY OF RELIGION

On the other hand, philosophical efforts concerning this domain are entirely of this world, and require a ceaseless struggle within this world to maintain their independence. The most essential weapon in this struggle is the purity and clarity of its own case, the more so since the philosophy of religion is often denied what is readily granted to less controversial disciplines — autonomy. Thus it is essential to introduce explicitly the postulate of an autonomous philosophy of religion, and to discuss its significance.

A modern application of the postulate of autonomy which points to its general character is the principle of the autonomous nature of science, which asserts the logical independence of science of any presupposition outside the field of knowledge (particularly of religious dogmas of all kinds). This principle was originally formulated in relation to the natural sciences, but generally speaking it holds equally well for philosophy and thus for the philosophy of religion. If the postulate of the autonomy of religious philosophy demands its logical independence of any presupposition outside knowledge, this is indeed a demand of a tall order. For the philosophy of religion (and similarly ethics and the philosophy of law) are far from complying with this demand).[2]

[2] Every attempt to comply with it is nipped in the bud by 'hermeneutic' reinterpretations of Leibniz's principle of sufficient reason, such as Heidegger for instance has advanced by means of reflection that 'hears through seeing, and sees through hearing'. In Heidegger's 'seeing ears' Leibniz's principle is then turned into the following 'existential proposition': 'Being and reason (*Grund*): identical; being: abyss (*Abgrund*)'; or else, coining an unusual verb: 'Being inheres in itself as its own reason.' Compared with this 'frenzy' of Heidegger's 'thinking' of the 'thoughtworthy', we must prefer Leibniz's 'calculating thought' — even though, in addition to being 'the discoverer of the principle of sufficient reason', he was also 'the inventor of life insurance'. (Cf. Heidegger, *Der Satz vom Grund*, Pfullingen 1957, pp. 89, 90, 105, 202, 211.)

Therefore (and for other reasons as well) it is necessary to examine the extent to which philosophical systems are autonomous and the extent to which they are not autonomous, i.e. authoritarian. A philosophical system can be directly or indirectly authoritarian. It is indirectly authoritarian when it sanctions other authoritarian systems, such as positive law and positive religion. It is directly authoritarian when it adopts authoritarian contents or methods. While indirectly authoritarian systems characterize the cultural environment from which they originate, directly authoritarian systems are the result of replacing methods of cognition by dictatorial edicts. The history of philosophy has produced a great number of pseudo-methods of cognition, which in practice replace argument with dictates, e.g. dialectics in the nineteenth century and *Wesensschau* (Scheler: the immediate grasp of essences) in the twentieth century. Both are surpassed by existentialism which elevates the dictum 'primum vivere, deinde philosophari' to a methodical axiom, thus ignoring reason not only in practice but also in principle. Heidegger's injunction to 'leap right into the circle' as 'the true measure of philosophic daring' — actually its transformation into metaphysical dogmatism — is typical of this kind of pseudo-philosophizing which, in league with neo-orthodox Protestant theology, has unfortunately acquired a markedly irrational influence in the field of religious philosophy.

On the other hand, rational (and thus autonomous) philosophy of religion has shown little inner development during the twentieth century and has retained little of its former, nor gained any new, influence. Its constructive rebeginning is a postulate that can only find fulfilment within the framework of a transcendental metaphysics which, based on the theory of the antinomies, knows how to draw on modern notions and experiences from such divergent fields as logic, astronomy and sociology.

Whatever the detailed results of such efforts, they will again represent the idea of the one and only religious truth (which can be reconciled with a multiplicity of forms of worship). But, over and beyond asserting this multiplicity, they will evaluate it just as they will evaluate all the forms of religious life itself (going as it does far beyond clerical cults). Because of the unspeakable and far too readily forgotten excesses of twentieth-century political religions, it will be of the utmost importance here to define clearly the limits of religious tolerance. The worship of blood-thirsty political gods, no matter how fanatical the devotion it inspires, must not be treated with tolerance, and a philosophy of religion which avoids practical conclusions of this kind, misses its highest vocation.

Turning its back on the idols of the day, the philosophy of religion looks at the universe *sub specie aeternitatis*, confident of the powers of reason, and thus in the spirit of Spinoza. It is not its function to undermine by sceptical analysis religious feelings, ideas and symbols and to reveal their alleged irrelevance. Its task, and a task which it alone can perform, is to be uncompromising with the demands of truth and their verification in the sphere of religion.

CONSCIENCE AND REASON

By Gustav Heckmann

I

What is conscience? How are we to understand what we experience within ourselves as the stirring of conscience? What do these experiences reveal to us about our nature, the nature of man?

These are the questions which will be discussed in this lecture.[1] It proceeds from the assumption that we cannot doubt the actual existence of the inner experiences which we call the stirring of conscience. We are dealing with a phenomenon which can be observed. We can perceive the stirring of conscience by introspection, and by accurate observation we are able to identify with some certainty a number of characteristic factors of the phenomenon.

First of all we will endeavour to ascertain these observable factors but will refrain from interpreting them. Only when we have done this will we attempt an interpretation.

The inner experiences known as stirrings of conscience include exhortation and warning.

The exhortation of conscience becomes audible to us in a concrete situation and expresses itself in some such way: 'Here you must do something. Here you cannot stand aside.' Sometimes it is quite clear to us what we have to do; at other times we find out only after a laborious search. Sometimes we feel at first merely uneasy and anxious and discover only later, often very much later, that a summons was trying to reach us and we were afraid of hearing it. In these forms we experience the exhortation of our conscience.

The warning of conscience makes itself heard when we intend to do something specific, and may be expressed thus: 'Stop! Don't do it! Wait! Think it over again!'

In both cases we are faced with a request, with a summons either to do something specific, or to refrain from carrying out a certain intention, or, again, to reflect on how we ought to behave in a given situation. This characteristic does not seem to me to be an interpretation, but a fact which can be observed.

The same characteristic can also be observed in the third form in which we experience a stirring of our conscience, namely in our subsequent judgment, although here the summons is, grammatically

[1] Enlarged text of the inaugural lecture given on his appointment as Principal of the Pedagogic Academy at Hanover on February 3rd, 1956.

speaking, expressed in the past tense. This judgment after the event may take some such form: 'How could you do such a thing! You should not have done that!' Or: 'You should not have kept quiet in this case! You should never have stood aside here!' It may also appear as a question implying some doubt: 'Was it right to act as you did?', and after examining our conscience honestly we may arrive at one or other of these replies: 'No, it was not right', or 'Yes, it was right'. In this case the hortative factor is included in the judgment that we have or have not disobeyed a summons.

Another factor which can be observed in the stirring of conscience and which, I feel, does not as yet include an interpretation is this: The summons contained in a stirring of conscience is always stimulated by and related to a concrete situation we experience. Moreover, the situation is always one in which our actions are interwoven with the well-being of others, the weal and woe of those with whom we have a reciprocal relation through our actions — our 'neighbours'.

The third factor that can be observed concerns our spontaneous reaction to a summons issued by our conscience — by spontaneous reaction I mean the natural reaction which has not been diverted by reflection. This is what happens: either we have no doubt about what we ought to do or are certain that we should not do a certain thing. It is then that we know: 'This is the right thing, this is correct', or 'That is wrong'. Or else we are in doubt as to what the right thing is, we are still in the dark about the right decision. Then we start searching for it. We examine the various possibilities, weigh reasons and counter-reasons. On occasions we seek the advice of a friend as to which decision is the right one. And we do not reject criticism of a decision we have taken on the grounds that such criticism is senseless from the outset. On the contrary, we examine it and either accept it or reject it as unwarranted and seek reasons for this rejection.

Our spontaneous reaction to the stirrings of conscience is therefore guided by the idea of that which is right objectively and is obligatory for us. It presupposes that the categories of right and wrong are applicable to our personal behaviour, to decisions prompted by the stirring of conscience. Actually in this respect we react no differently to the stirring of conscience than we do to any set of circumstances which have to be weighed objectively; our spontaneous reaction, undisturbed by philosophical doubts, has that characteristic. It proceeds from the assumption that the idea of truth can be applied to the decisions prompted by the stirring of conscience.

This also appears to me to be a fact that can be verified by introspection. We must, however, watch carefully how far observation

takes us. Introspection shows that in actual fact we accept the stirrings of conscience as something that is subject to the idea of truth. To decide whether this assumption is justified or whether it is a prejudice, goes beyond the mere observation of this phenomenon. Here, it seems to me, we face the most important problem of the interpretation of the observable facts: what is the meaning of the claim to truth of the stirring of conscience, a claim which we accept by our spontaneous behaviour?

Exception might be taken to our assertion that spontaneous acceptance of its claim to truth is one of the observable factors of conscience. It might be pointed out that our spontaneous reaction is sometimes quite different in that it represses and stifles the voice of conscience. However, this does not alter the fact that as soon as we allow ourselves to become conscious of the stirring of conscience we accept its claim to truth. For what do we fear when we suppress it if not this claim? We feel it even while we suppress it.

Let me continue enumerating the factors that can be observed in the phenomenon called conscience. The fourth concerns the question of where we seek the reasons for our decision when, in conformity with the claim to truth of the stirring of conscience, we try to find a morally justifiable decision. We look for them in the concrete situation of reciprocal relationships to our neighbours, or more precisely in the effect of our decision on the weal and woe of others.

The fifth factor is the frequent uncertainty, the vacillation of our moral decisions, the frequent inability to agree with others as to the right moral decision and the dependence of man's moral convictions on his cultural surroundings.

It is this factor which is especially emphasized by the supporters of ethical relativism when they deny the applicability of the idea of truth to our moral decisions. The compatibility of this factor with the spontaneously accepted claim to truth is therefore an essential problem in the interpretation of the facts verified by observation.

We will now turn to the questions of interpretation.[2]

II

The interpretation which I shall now present for your consideration has as its basis Kant's recognition of the co-operation of sense

[2] For the observation of this phenomenon see the chapter on Descriptive Analysis of Moral Feeling in *Vorlesungen über die Grundlagen der Ethik, 1. Band, Kritik der praktischen Vernunft,* by Leonard Nelson (Leipzig 1917). Nelson is particularly concerned with methodical clarity in the foundations of ethics. He makes a precise distinction between the different steps: search for the assumptions underlying our moral judgment, observation of the facts concerning our moral feeling, and interpretation of these facts, and he examines the importance of these different methodical steps.

perception and reason in our rational life. I should like to illustrate this co-operation by some examples first from the sphere of perception, then from that of desire and will. For this I ask your indulgence. I hope that it will be rewarded by the proof that conscience fits naturally into the whole of our sensual-rational life.

It is possible to observe the co-operation of these two forces even in the simplest cases of every-day experience. When I look out of the window and say 'I see snow', this assertion contains much more than what is conveyed to me by my vision at the moment of looking. What I really see at that moment, i.e. perceive through the sense of sight, is simply an expanse of whiteness with glitter and shade. No more is conveyed by actual perception, and perception always depends on the presence of an impression. The statement, 'I see snow', contains much more. By my interpreting the present sense-impression as snow, I am classifying it together with an untold number of earlier sense-impressions as an experience. I now comprehend the whiteness conveyed to me as a sense-impression as a white entity, as the same substance whose whiteness I have perceived on previous occasions and of whose other properties I have been made aware in other instances by other senses than the sense of sight, properties such as the slushiness which I have felt with my hands, the coldness which I have felt on my forehead, the grating which I have felt and heard when I crushed the snow, the insipid, watery taste on my lips. All this is not conveyed to me in the present sense-impression of whiteness and therefore the statement, 'I see snow', goes far beyond it. It brings together a large number of past sense-perceptions with the present one, forming a single empirical cognition.

This bringing together of sense-impressions registered at different times and through different senses becomes possible only because a concept is added which is not contained in any one of the combined perceptions: the concept of property-of-a-thing. We comprehend the whiteness conveyed by perception as the property of a thing, a substance, the same substance whose other properties we have become aware of by previous perceptions. If we pay sufficient attention to that which was conveyed at the time by each of the perceptions which contributed to the formation of so simple an experience we realize that the concept property-of-a-thing was not included. By reflection we add this concept to the properties we have perceived and only thereby do we co-ordinate our incoherent sense-perceptions to a complete cognition in which we comprehend white, cold, slushy and grating as properties of one and the same substance.

The concept property-of-a-thing is not the only co-ordinating,

purely reflective concept which enters into the statement, 'I see snow'. Another is the concept 'identical'. The substance of the white stuff is identical with that of which I noticed those other properties on previous occasions. However, it is not important here to proceed further in our analysis of experience. The point at issue is this: even the simplest experience is possible only when the faculty of perception co-operates with a faculty of abstract concepts, concepts of an entirely different kind from those conveyed by perception. With Kant, we call this faculty reason.[3]

I would like to dwell a little longer on the example of the snow in order to show also that the concept of cause and effect already plays a part in it. This becomes clear when I formulate my statement somewhat differently: 'I see that snow has fallen.' In the strict sense of perception I can certainly not see this. For perception is tied to the visible present and a past happening cannot be perceived.

The perception of whiteness, however, incites a series of associations in my mind: I compare the present impression of whiteness with another impression I can still remember, the impression I received of the same garden last night, and I interpret the change I notice as the result of a snowfall which happened during the night. This interpretation is only possible through the assimilation of a large number of previous perceptions which I have had when it was snowing. It contains more than one abstract concept which is not present in any perception: first the concept of happening, of the event, of reality existent in time, and then that of cause and effect, the concept of the actual connection between events following one another in time.

To sum up: If reason did not play a part in our cognition, it would not be possible for us to have even the simplest experience. The theory of radical Empiricism, according to which perception is our only faculty of cognition, cannot explain how even the simplest experiences are brought about. The critical antithesis — in the sense of critique of reason — is as follows: even the simplest of experiences include co-ordinating concepts which are not contained in any perception and therefore must have their origin in a different capacity of the mind from that of perception.

A few observations on the part played by the two co-operating faculties. The faculty of perception plays a receptive, passive part in our cognition, reason an assimilative, formative and active part. The passive element in perception lies in the given fact of the sense-

[3] For the sake of simplifying the argument I shall not discuss the part played by the ideas of space and time, though they, too, are to be found even in the simplest forms of experience. It is sufficient here to clarify the difference between the faculty of perception and that of reason and to consider the co-operation of these two faculties.

impression which we can only register as it is conveyed. The active element, or at least the comparatively active element, in reason is to be found in the fact that we add the co-ordination of the elements which are incoherent when conveyed by perception, only *after* we have received the sense-impressions. We realize that we actually do this when we have erred in the process of co-ordination and rectify the error afterwards. We believe, for instance, that we see two poplars in the distance but on coming nearer we find that they are the spires of a church. We can also observe how a baby has first to learn to co-ordinate sense-impressions before it can form complete cognitions; we can observe how it is occupied, for example, with the problem of combining the visual impression of its foot with the impression of feeling the same foot when it touches it with its hand, and with the experience that the foot is attached to its body, in forming the whole of a cognition. At first the infant has only a chaos of sense-impressions, yet although we ourselves must have lived through this state, it can only become intelligible to us through a process of abstraction.

All the so-called sense-delusions are due to erroneous co-ordination. It is not the senses which delude us; their impressions as given facts cannot be doubted. The error lies in a false co-ordination of the sense-impressions. We correct such an error by virtue of further sense-impressions, by extending the range of observations within which we co-ordinate them.

Jakob Friedrich Fries described reason as spontaneity which can be stimulated. It is only when sense-impressions are conveyed that reason becomes active, that it is stimulated to its quest for co-ordinating the single, incoherent elements to form a harmonious whole. I should like to sum up the part played by reason in the words of Paul Bernays:

> In natural science we find that our reason is not immediately cognizant of the basic foundations of its explorations, but is constantly searching for them. Its productivity manifests itself in the formation of interpretative conceptions in which we classify what approaches us first through perception and then through more complex experience; in this way a development of conceptual interpretation takes place, and through this development, a store of interpretative categories is accumulated in the 'collective unconscious', to use a term of modern analytical psychology.[4]

[4] See 'Ueber die Fries'sche Annahme einer Wiederbeobachtung der unmittelbaren Erkenntnis', p. 128, in *Leonard Nelson zum Gedächtnis*, edited by Minna Specht and Willi Eichler, Frankfurt-am-Main and Göttingen (1953).

I should like to emphasize two ideas which to me seem to govern the activity of reason. The one idea is that of truth. In the simple example of the snow this idea appears in its most modest beginning. We seek for the correct co-ordination of the sense-impressions. We may recognize that we have made a mistake in this co-ordination and correct it. We form criteria for right and wrong, for truth and error. To me it appears that the idea of truth penetrates all our rational activity, even in the regions of desire, evaluation and will, which we shall consider in a moment.

The other idea, that of the whole, which seems to me to be closely linked with the idea of truth, also governs all our rational activity. To start with, it is a matter of advancing from disconnected sense-impressions to a complete cognition in its simplest form. As our experience extends further, errors are recognized as such and corrected. This is done by taking into account further perceptions and experiences, by extending the scope of our isolated sense-impressions and forming them into a harmonious whole. This harmonization serves as a criterion of truth.

From the elementary stage in which we first saw them at work these two ideas, that of truth and that of the whole, press forward in our mind, trying to transgress every possible frontier, and thus lead us towards forming our deepest convictions.

III

We now turn to the regions of desire, aspiration and will. Is it the case that here, too, co-ordination between sensibility and reason can be observed? Is there, in this respect, an analogy between this region and that of experience and cognition? This question provides a fruitful, heuristic starting point for our discussion even though we cannot take it for granted that our observations in the one sphere necessarily apply to the other. We must find out whether there is an analogy and how far it extends.

First of all we notice that the lack of coherence which we must have experienced in the sphere of cognition in our infancy when we were only able to register chaotic disconnected sense-impressions, to some extent still holds for the region of the will. Take, for example, the mental condition described by the phrase that someone has lost his head, or is out of his senses. We experience this mental condition, for instance, when we are in a state of anxiety, perhaps under the pressure of too many and too heavy tasks with which we are unable to cope. Then the following may happen: We apply ourselves to the first thing that comes along but turn from it under the impact of a

fortuitous association or impulse in order to embark on another task, which again, driven by a new chance impulse, we allow to slip away. Thus we stumble from one thing to another, without any plan, one moment doing this and another moment that, without achieving anything. When we are in this state our will is a succession of disconnected impulses. It lacks an element which would create a co-ordination of impulses over a period of time. This condition can be compared with that of chaotic, disconnected sense-impressions, which exists in early childhood in the region of cognition.

What is the element necessary here to achieve co-ordination which is missing in the condition of losing one's head? We can discover it by calling to mind how we extricate ourselves from such a mental condition. We start considering what is really important now. In asking ourselves what is the most important thing to do and in answering this question by our insight into that which is essential, the element which brings about co-ordination comes into the picture. As soon as we are guided by such insight our will achieves unity of action over a period of time and is no longer a succession of disconnected impulses.

Let us compare the element which brings about co-ordination with the elements which were previously disconnected, in other words the insight into that which is essential with the impulses springing from an anxious, aimless restlessness whose objectives are determined by fortuitous associations. Thus we recognize the one as being a part of rational activity, and the other as belonging to the sphere of sensibility, corresponding to the characteristic difference between the two faculties of the mind to which we referred in the region of experience and cognition.

The state of being pulled to and fro by successive impulses is a given fact just as a sense-impression stamped upon us is a given fact. We find ourselves in this state without seeking it. The individual impulses, too, appear and stir us without any encouragement on our part. This is the element of passivity which we have found characteristic of the senses. That state is a state of shapelessness; what our senses have conveyed has not yet been formed into a unity of action. This further development only occurs with the aid of a new concept and a new impulse which were not present in that which was given without our seeking: the idea of what is essential in a given situation and the insight into that which is essential, which acts as impulse. Here we see reason acting in the same way as it does in the region of cognition. Having been stimulated by incoherent sense-impressions reason makes its own contribution by adding concepts and impulses that bring about co-ordination, and with their aid develops

our will into a harmonious whole. We also find here at work the other idea which governs rational activity, that of truth. We seek reasons to test our decision as to that which is essential and are prepared, if necessary, to correct this decision.

Another example of incoherence of willing is the case of the person who is pulled to and fro by changing moods and passions, who – as we say – is at their mercy. The words which describe this condition clearly show the element of passivity which is characteristic of the senses. As in our previous example, we can discover the missing element of reason by asking how we can free ourselves from this mental condition. This seems possible only by insight won from painful experience; by learning to understand that it is better to control our passions. Better for what? For our health, for a happy life, for the welfare of our family, for professional success, and many other things.

Thus, our understanding and our aims, through which we achieve unity of action in our will, may be narrow or broad. Our aims may be narrowly circumscribed by what we believe will be advantageous for ourselves or we may be concerned with a wide range of values. But even action guided by the most utilitarian considerations contains an element of reason. In so far as action is at all governed by an aim, the condition of that which is merely given, of mere impressions stamped upon us by external and internal conditions, is already overcome. The ideas of truth and of the whole, which govern our rational activity, are already at work here in their simplest form: the idea of truth in the form of a judgment of the type, 'It is better', even if the 'better' is determined by a narrow point of view, and the idea of the whole in the form of unity of action over a period of time.

In these examples we have considered the disconnected willing in time, in the chronological succession of impulses. It also appears in quite a different form in the case of a split personality. We refer here to the phenomenon of the behaviour of a person being governed by motives which are completely unconnected. His mode of behaviour on one occasion is so completely different from that on another occasion that we wonder how the same person can behave in such different ways. If we have known him previously from one side only and suddenly discover his other self we may well ask ourselves, 'Is this the same man?'

Such a person has failed to develop some sort of unity out of the motive-forces within himself. Perhaps he has not consciously assigned to the elementary impulses a place in his willing and has not brought value-convictions and elementary impulses face to face with one another and assimilated them to a single willing. This task

which he has failed to perform is one that can only be undertaken by reason, and only the ideas of truth and of the whole can help to find a solution. How much insight into truth is required, how much understanding of our own nature, how much deepening of our value-convictions so that they pervade not only our intellect but our whole being — all this is convincingly shown by the experience of psychotherapy. For the task of psychotherapy is the integration of a man's impulses, convictions and philosophy of life into a complete person.

The discussion of our examples of disconnected willing suggests to us that the impulse of reason which urges unity of action ever tries to push beyond all frontiers that are recognized as such, and to advance beyond all partial aims that are recognized as such. For the question, 'What is better?', leads us inevitably beyond mere considerations of personal advantage and beyond more limited evaluations to the question, 'What gives value to life as a whole?'

Goethe, at the age of thirty, described the impulse of reason which urges us to be concerned about the whole of our life in these words:

> The day-to-day tasks which I accept as my duty become both easier and more difficult every day. They claim my attention in my working and in my dreaming existence. Every day this duty becomes dearer to me and in it, and in nothing greater, I wish to equal the greatest of men. This desire to raise up into the air as high as possible the pyramid of my existence whose basis has been given to me outweighs everything else and hardly admits even momentary forgetting. I must not delay, I am already far advanced in years and, perhaps, fate will break me in the middle and the Tower of Babylon will remain stunted and incomplete. At least it should be said that it was boldly designed, and should I live, my strength shall, with God's help, last till the top is reached.[5]

IV

Yet with this, the path of reason which leads towards an ever-increasing integration of separate elements into a whole has not come to an end. The important step that leads further than we have gone so far and that conscience requires of us, still remains to be taken. The meaning of its summons appears to me to complete the path from the Ego to the Thou.

We have not yet spoken of this path in actual words. Those who

[5] Letter to Lavater from Ostheim vor der Rhön, about September 20th, 1780. Memorial edition of his works, Artemis Publ., Zürich: Letters of the years 1764-86, p. 532.

try to live their lives in the spirit of the quotation from Goethe have in fact embarked on this path. In so doing, they constantly rely on others, on community, on exchange of ideas, on being understood, on friendship and love. The need of community is most evident in the child, and from the very first day of its life. For the child, to feel secure with those around him, above all with its mother, is an elementary psychical need. And for adults the give and take in love relationships with others remains a vital necessity in life. To this extent the striving for community, the path from the Ego to the Thou, is already contained in our striving to achieve completeness in our own life.

Within this impulse of reason, however, the Thou is seen only from the angle of our own life. So long as we are entirely absorbed by the desire to raise the pyramid of our existence as high as possible, the Thou is seen from the point of view of life-fulfilment of the Ego. Seen from the side of the other person this point of view is narrow and naive. All of us, especially the men among us, know how we can be caught up in this narrowness. He who is possessed by his task in life easily succumbs to it. He devotes all his efforts to building his pyramid and suddenly finds that he has never seen the other person, whom he constantly needed, as a separate being with his own interests.

At this point conscience summons us to a new important step, a step of reason in the direction of the whole. We are challenged by the question: Can you who rely on love, you who receive love and give love, stop short at a relationship to the Thou which is entirely governed by the idea of your own life? Can you be satisfied with that, knowing that the other person is a human being too, destined to raise the pyramid of his existence as high as possible? To accept the Thou for its own sake, and not merely from the point of view of the interests of the Ego, no matter how noble the aim pursued by it, seems to me the meaning of the summons issued by our conscience.

But are we not already reaching beyond the narrow Ego-governed relationship to the Thou by the love relations which are essential to our own life? It is true, love is in full accord with the demands of conscience. And the person who is rich in love and whose relations to others are largely permeated by this power will be more rarely questioned by his conscience than one who lacks human warmth. He acts from spontaneous kindness.

And yet, even the kindest of persons will sometimes become so absorbed by the problems of his own life purpose that he sees the other person only from the point of view of his own problems. In such a situation an alert conscience will raise its voice.

There is yet another situation in which a kind person especially will wrestle with his conscience: when he feels called upon to bring danger and grief upon those dearest to him for the sake of a higher duty. The men of the resistance to National Socialism were confronted by this question.

This example shows clearly how much may be asked of us by our conscience when it summons us to overcome our self-centredness. When our conscience asks us to accept the Thou for its own sake and to act according to our recognition that the Thou and the Ego are beings of the same kind with equal rights, then it asks us not to limit this step to those dearest to us. For this would merely be a new form of self-centredness, a refusal to go beyond the interests of the group of people to whom we are bound by close personal ties.

The impulse of reason which acts within us under the impact of the ideas of truth and the whole, transgresses all group limitations. Untold grief has been caused by raising the interests of a group to the status of the supreme moral law. Our 'neighbours' are all who are affected by our actions, and in modern society this is a large and ever-expanding circle.

We might say that this is how our conscience speaks to us: because as a human being you are in need of community, you have to accept it for what it is. It is not made to suit your interests; in it you are not more than your neighbours, but their equal. Accept this law as basic to community. Take this further step from separateness towards the whole, going as far as the conception of who is your neighbour will carry you.

V

In conclusion let us consider the question: What does our interpretation of the phenomenon of conscience accomplish?

It respects the law of continuity in our mental development. It understands the stirrings of conscience as a link in a continuous sequence of mental activities which, through the co-operation of sensibility and reason, begins with simple acts of cognition and willing. It makes the development of that sequence understandable by virtue of an all-pervasive impulse of reason, and the links in the series, from the most primitive to the most developed, as expressions of the same impulse of reason.

In this connection a further argument in support of our interpretation may be mentioned here: its compatibility with the continuity of development in nature from earlier forms of animal life to the human species. A conception of the reasonable nature of man, if it is to be satisfactory from the scientific point of view, must accord

with the results of natural science, particularly with the Theory of Evolution. When Kant developed the Critique of Reason, this question had not yet been posed, but today we have to make this requirement of the Critique of Reason. Our interpretation of conscience satisfies it. For the co-operation of sensibility and reason, as we have described it, or — more briefly — experience, can also be found in its beginnings in animals. What distinguishes man from the animal appears to me to be the higher development of experience which has become possible through speech and through abstract thinking. We have assigned to conscience a place in a continuous series of ever-advancing experience. In so far as the Theory of Evolution can explain the development of new kinds of psycho-mental ability, such as speech and abstract thinking, the evolution of conscience becomes comprehensible.

Finally, it appears to me that our interpretation of conscience is in accord with the facts of introspection as we have expounded them above when describing the five observable factors.

The first of these factors was the hortative character of the stirring of conscience. We have interpreted conscience as a call to reason to become active in transcending egotism, even ennobled egotism, and to advance towards the whole. It lies in the nature of the step which conscience asks of reason, that it appears to us not as a spontaneous urge within ourselves, but as a severe, distant summons.

The second factor we observed in the stirring of conscience was the part played by the situation we experience, which is always one of interaction with others. Reason is spontaneity which can be stimulated; it becomes active only by virtue of the stimulus given by sense-impressions in a concrete situation. That the step taken by reason which conscience demands is prompted by a situation in which we are in interaction with others, is quite understandable again by reason of the special nature of this step.

There is no difficulty about our interpretation of the third factor observed by us: our spontaneous acceptance of the stirring of conscience as an expression of our capacity of truth. Our interpretation states: The stirring of conscience is an expression of our capacity of truth. It is an expression of our capacity of truth in the same way as every other rational activity, every co-ordination of separate sense-impressions to a congruous, harmonious whole. It is this inner harmony of the whole that constitutes the criterion of truth.

This factor, which is easily understood by virtue of our interpretation, creates for all interpretations of the phenomenon of

conscience, which differ fundamentally from ours, insurmountable difficulties. By interpretations which differ fundamentally from ours I mean those which deny the applicability of the idea of truth to questions and hints which come to us through our conscience. There are chiefly two such interpretations. Freud interprets conscience as the result of drive-force destinies (Triebschicksale) in childhood. Drive is understood here as a basic given fact — of a sensual kind in our terminology — and not in the sense in which we have spoken of the impulse of reason which presses forward towards an ever more comprehensive co-ordination. The other interpretation is sociological in character: it sees in conscience a product of social influences, in particular of education.

Doubtless, education and social influences strongly affect the development of our conscience. In the terms of our interpretation: how far the innate capacity of truth which rouses our conscience becomes a pure and powerful force, how far it remains confined and narrow, how far it is diverted and distorted, how far it becomes stifled — all this depends on diverse social influences. It also depends on 'drive-force destinies', that is on whether the energy of elementary urges (Triebenergien) goes more into the channels of love than of hate, more into the channels of composed self-reliance than of self-contempt, more into the channels of fear than of calm security. But it presupposes an original faculty of reason, and that means a faculty guided by the ideas of truth and the whole, which either develops or is crippled.

The fundamentally different interpretations deny precisely this supposition. The spontaneous acceptance of the claims of conscience as claims of truth, which we observe as a fact, is regarded by these interpretations as mere prejudice. But they have then to explain this fact psychologically, that is, on the basis of their own supposition. They have to answer the question: How can it be psychologically explained that the idea of truth, which according to their own supposition can have no meaning outside experience or the cognition of being, is transferred to a region in which it is — allegedly — entirely inapplicable, the region of moral decision? Furthermore, how can it be explained that in our natural, spontaneous reaction this transfer invariably happens? How can it be explained that only secondary, theoretical and never spontaneous arguments may temporarily arouse doubts about this? This question must be answered by every theory which denies the claim of our conscience to truth. It is not sufficient for such a theory to point to the many aberrations from the path of truth to which our conscience is undoubtedly exposed under the impact of elementary urges or

social influences. It appears to me that when put to the test of that question all theories denying the claim to truth of conscience fail.[6]

This means that there is moral truth and that conscience is a source of truth despite all the difficulties and dangers to which the comprehension of truth is exposed particularly in this region.

The fourth observable factor, namely that in our quest for a morally right decision we seek the reasons for our decision in the concrete situation which brings us into a reciprocal relationship with others, is not difficult to understand in the light of our interpretation. What our conscience asks of us is simply that we should look at the situation in such a way as to assign no place of privilege to the Ego over the Thou.

There remains the fifth factor, the factor which is invoked particularly by the ethical Relativists: the vacillations in our moral convictions and judgments, and the differences of opinion which we encounter here, especially amongst people of different civilizations. In the face of these uncertainties, how are we to understand that truth exists?

Let us face the fact that truth is often difficult to recognize and anything but evident. And the general truth which announces itself through the voice of conscience is of that kind. In a concrete individual case, our conscience may speak with a certainty and power of conviction which excludes all doubt as to what we have to do in this particular situation. But the general truth to which our conscience draws attention is difficult to comprehend and is not evident. Our interpretation makes it clear why this is the case. For what is it that conscience requires of us? We are called upon to free ourselves completely from the preference we give to the interests of the Ego in our thinking and willing if we follow our natural inclinations. We are called upon to be so deeply aware of all our neighbours and their interests that the Ego with its interests has no greater pull than the Thou with its interests. This general truth of which conscience speaks to us is necessarily abstract. For it must leave room for the unforeseeable in any situation; for new factors and new problems which may appear in the process of social and human development; for the difference in the problems of different nations and civilizations. This general truth cannot be expressed in commands or prohibitions of the type: Thou shalt not lie; thou shalt not steal. Situations may exist in which our conscience asks us to do something which generally speaking we would consider wrong. The general truth — if we are to formulate it at all — can therefore only lay down the method to be used in seeking honestly and taking

[6] That would have to be shown by a thorough discussion of these theories.

account of all the relevant circumstances, the morally correct decision appropriate to the whole of the actual situation.[7] But here the question arises: What sense is there in formulating this general moral truth? If in a particular case we wrestle with our conscience, then we are not influenced by the general moral truth in its abstract form. It is by our feeling for truth that we are moved to turn our attention unreservedly to the actual situation and to find out how the general moral truth may be adequately applied to it.

In view of the nature of this general truth the manifold difficulties which confront the attempt to clarify the conception and to obtain agreement on it can easily be understood. Our interpretation of conscience, therefore, is also in accordance with our fifth and last observable factor.

I see no reason, therefore, which would compel us to admit an antagonism between our scientific-critical attitude and our spontaneous natural reaction which accepts conscience as a source of truth. On the contrary, I find that our interpretation of conscience which corresponds to this natural behaviour does more justice to the observable facts than those interpretations which deny the claim of conscience to truth. I find also that to understand the idea of truth as an idea governing not only our cognition of being but also our decisions, fits harmoniously into the whole picture of the sensual-rational character of human beings.

[7] Nelson defined this state of affairs as the formal character of the moral law and discussed it in many places in his 'Critique of Practical Reason'. The relevant passages can be found listed in the subject matter index. Compare also Nelson's 'System of Ethics', Yale University Press, New Haven (1956), pp. 91-6.

A MODERN SAMSON IN THE PHILOSOPHER'S TEMPLE: MORTON WHITE'S NEW BOOK

By J. W. N. Watkins

Modern philosophy is supported by two main pillars: the distinction between analytic and synthetic statements; and the distinction between descriptive statements and prescriptive utterances. Professor Morton G. White, in his new book,[1] has succeeded — to his own satisfaction, at any rate — in pulling these two pillars down. Unlike the original Samson (and unlike the Wittgenstein of the *Tractatus*) he escapes the effects of his work of destruction, and remains cheerfully alive to offer brisk and egregious advice about the disposal of the ruins.

It is a bold book, and one with a strong sense of intellectual movement. It transports us from Cambridge, England, before the First World War, via inter-war Vienna, to Cambridge, Massachusetts, in the 1950s — in other words, from what White calls analytic platonism, via logical positivism, to the most up-to-date transatlantic neopragmatism. At the outset, the number of animals in the philosopher's zoo is admittedly large, but at any rate each animal is in its own cage. By the end, all the rarer and more peculiar animals have been killed off; the remaining animals have been released from their cages and are busily inter-breeding.

The book is extremely well organized. It is divided vertically into three main parts: 'What Is', 'What Must Be', and 'What Should Be'. Within each part a parallel development is unfolded which divides horizontally into three main stages. The first is the ontology of Bertrand Russell and G. E. Moore before 1914. This included timeless universals, relations, physical objects, synthetic *a priori* truths, and simple and non-natural qualities. Next comes the early positivist stage when it was claimed that sentences in natural languages are correctly *described* as analytic, empirical or meaningless, according to their inherent characteristics. Difficulties here lead to the later, more conventionalist stage (which White calls 'the age of decision') when the positivist philosopher declares that he intends to *call* these sentences 'analytic' and those 'meaningless'. This leads White to the contention — he acknowledges a considerable debt to Professor Austin[2] — that philosophers' utterances are mainly performative ('I

[1] *Toward Reunion in Philosophy*, Harvard University Press, 1956. Pp. xv+308.

[2] Professor J. L. Austin has broadcast and frequently lectured on the topic of performative utterances but he has not, so far as I know, published a systematic account of his

postulate . . .', 'I define . . .', etc.) and that their problem is how to *justify* such performances. Since he believes that a philosophical decision is never dictated by 'facts', White concludes that its justification is a quasi-moral affair. Whereas moral philosophers usually feel that their subject is in a less satisfactory state than, say, the philosophy of science or of law, and often turn to these for illumination, White feels that the difficulties of 'the morals of philosophical legislation' oblige him to turn to ethics for illumination, even though he believes that moral philosophy has been infected by errors similar to those whose detection and removal in Parts I and II of his book led him to his quasi-ethical conception of philosophical activity.

In Part IV he assembles the consequences of his earlier criticisms, eliminations and decisions. These consequences are startling, to put it mildly, and an account of them may be postponed until the reader has had more time to prepare himself. I will now set out what I take to be the main lines of White's argument in the first three parts of his book. I shall venture a few incidental criticisms; but my main criticisms will be reserved until his whole argument and its implications are before us.

Philosophers' Entities

Part I discusses the status of various philosophical entities — relations, physical objects, things-in-themselves, etc. — but, from the point of view of what is to follow, far and away the most important discussion concerns *meanings* considered as non-physical, non-psychical entities. Before considering White's reasons for repudiating the existence of such entities, however, it will be as well to mention briefly his views about the status of philosophers' existential statements in general.

Philosophers assert and deny the existence, not so much of specific things like the Tower of London, but of whole categories of things,

views on it. Very briefly, he has driven home the point that a number of utterances (usually in the first person singular, present tense) which look like, and have often been regarded as, descriptive assertions, are not really descriptions of anything, but simply acts or performances. Thus whereas 'I promised', 'He closed the meeting', or 'The name of this ship is *Mayflower*' are indeed descriptive statements, 'I promise', 'The meeting is closed' (announced by the chairman), and 'I name this ship *Mayflower*' are what Austin calls performative utterances. In his 'Other Minds' (*Logic and Language*, II, ed. A. G. N. Flew, especially pp. 142-7) Austin pointed to an analogy between 'I feel sure' and 'I fully intend' on the one hand, and between 'I know' and 'I promise' on the other. 'Saying "I know" ', he writes, 'is taking a new plunge. But it is *not* saying "I have performed a specially striking feat of cognition, superior, in the same scale as believing and being sure, even to being merely quite sure": for there *is* nothing in that scale superior to being quite sure. Just as promising is not something superior, in the same scale as hoping and intending, even to merely fully intending: for there *is* nothing in that scale superior to fully intending.' He puts 'I know' in the same boat with 'I swear' and 'I guarantee': to utter one of these formulas is to commit yourself, to stake your *reputation*, to assume a special responsibility for what you say.

like minds, bodies and numbers. This has given rise to two views, both of which White rejects. The first is Professor Ryle's view that the meaning of 'there are' varies with the category of the predicate it quantifies. Thus it means one thing in 'There are gloves', another in 'There are pairs', and another in 'There are minds'. White maintains to the contrary that the existential quantifier of modern logic has a univocal meaning. The most he will concede is that some puzzling existential statements may be translated into a non-existential form which removes the puzzlement. The second is Professor Carnap's view that ontological existential statements like 'There are physical objects', as opposed to empirical existential statements like 'There are cabbages', do not describe the world but lay down a linguistic framework, a 'universe of discourse', within which the world can be described. They are linguistic stipulations rather than factual statements, useful or useless rather than true or false. Against this White insists that no predicate is *inherently* a category-predicate (it is a major thesis of his book that no word or sentence is *as a matter of fact* undefinable, meaningless, analytic, etc. — we make it so). We may take physical objects as our universe of discourse within which 'cabbages exist' is a synthetic statement. Equally, we may decide, White says, to 'take cabbages as our universe of discourse'; will 'cabbages exist' thereby cease to be true or false? No, for this decision could not be justified if 'cabbages exist' were not true. Moreover, the truth of an existential category-statement is entailed by the truth of a statement asserting the existence of something within the category. The fact that an existential statement is highly general, even the fact that it is as general as it could be within some universe of discourse, makes it no less synthetic than a more specific existential statement.

White next considers the early positivist contention that ontological existential sentences are, like other unscientific and non-analytic sentences, meaningless. He easily shows the artificiality of the claim that certain sentences are inherently meaningless:

> What happens when a criterion is presented according to which decent scientific sentences are meaningless? The criterion is revised. What happens when atrocious metaphysical sentences turn out to be meaningful? The criterion is revised. But how do we know whether the offending sentence *is* scientific or metaphysical? This is never made clear. . . .

Since sentences are not inherently meaningful or meaningless the most one can do, White says, is to *call* this type of sentence 'meaningful' and that 'meaningless'. This will be a performative utterance

rather than a descriptive statement. The big question now becomes, Can such performances be justified? In White's eyes, the main effect of Austin's work is to show that philosophy is, at bottom, a quasi-ethical activity.

With this outcome of his discussion of existential statements in philosophy before us, we can now consider his answer to the particular question, Can the decision to postulate *meanings* be justified?

Perhaps the simplest and most striking argument in support of the metaphysical view that as well as words and the objects they denote there is a further realm of non-physical, non-psychical meanings, was given by Frege when he pointed out that 'the evening star' does not *mean* the same as 'the morning star' although they both denote the same thing. Surely one can only distinguish the *a priori* character of 'the morning star is the same as the morning star' from the *a posteriori* character of 'the morning star is the same as the evening star' if one recognizes the *meanings* of these two terms? White goes on to make the important historical point that in 1912, after Russell and Whitehead had succeeded in deducing arithmetic from purely logical notions, Russell in his *Problems of Philosophy* still regarded arithmetical truths as synthetic *a priori*. Admittedly, this was partly dictated by purely verbal considerations: Russell accepted Kant's definition of an analytic truth as one whose predicate is contained in its subject, and this definition does not fit arithmetical truths. But much more important was the fact that Russell believed in the existence of a timeless world of meanings, and he regarded mathematics as a set of factual assertions about the necessary relations between these meanings.

But White is utterly unimpressed by the argument that if we are to *understand* a word (as opposed to merely knowing what it denotes) and more especially, if we are to recognize the *a priori* character of logical and mathematical truths, there must exist meanings we can grasp and necessary relations between them which we can perceive. White rejects the view that any statement is inherently *a priori* — '*a priori*' statements are simply those we have *resolved* never to give up. And he claims that the fact that people understand words no more entails the existence of meanings than the fact that people worship entails the existence of God. There are, he holds, no factual or logical considerations which prove the existence of meanings; there is, on the contrary, a strong pragmatic reason for refusing to postulate them. The reason he advances is actually twofold, though he appears to regard it as a single reason: first, 'meanings' are mysterious, occult entities; secondly, they have been postulated in a purely *ad hoc* way to explain facts which either need no explanation or need a much better

explanation. There is no evidence, he says, for their existence other than the evidence they were invented to explain. They are otiose hypostatizations, like the 'dormitive virtue' in opium. This is the main conclusion of Part I, and it is wielded in a pretty drastic way in Parts II and III to undermine the distinction between analytic and synthetic statements, and the distinction between natural and non-natural predicates. Later, I shall argue that the large and dramatic consequences which follow the repudiation of meanings rebut the main claim on which their repudiation was founded, namely, that they are merely *ad hoc* inventions. But let us first see what these consequences are.

Analyticity

White begins by establishing the important point that if an *a priori* truth is a statement which can be known to be true without recourse to experience then, although we can say that all analytic truths are *a priori*, it is problematical whether all *a priori* truths are analytic. I should have liked him to have gone still further in separating these two notions. I would call *a posteriori* a statement which can be tested by experience, and *a priori* a statement which can *not* be tested by experience. This would allow the possibility of at least three kinds of *a priori* statement: those which are synthetic and necessarily true, those which are synthetic but not necessarily true, and those which are analytic. White recognizes the first and third possibilities, but does not mention the second. I believe that most important metaphysical doctrines are of the second kind. But White, like most critics of the analytic/synthetic dichotomy, believes that it is the notion of analyticity, rather than the notion of syntheticity, which causes all the trouble.

He begins his critical examination of analyticity by dividing so-called 'analytic' truths into two classes: the first comprises logical truths, or exemplifications of such logical laws as 'Every P is a P' and 'Every P and Q is a P'. For the believer in analyticity the only problem raised by such statements as 'Every female fox is a fox' is the problem of showing that the laws of logic are analytic. The second and much more problematical class comprises statements such as 'Every vixen is a female fox' and '$7+5 = 12$', statements which are not *obvious* exemplifications of logical laws, and to demonstrate the analyticity of which certain linguistic operations would have to be performed to translate them into logical truths. White denies the legitimacy of such operations.

In order, he says, to prove the analyticity of mathematical truths you need 'a logic which (*a*) sets forth its general laws, (*b*) shows them

to be analytic . . . , (*c*) shows them to be sufficient for the deduction of all *a priori* knowledge, and (*d*) seriously meets the demand of showing that all so-called *a priori* truths are analytic in this way'. In support of his denial that such a logic exists he quotes Russell's admission that 'the assertion that the number of individuals in the universe is *n*, whatever number *n* may be' is not a tautology, yet it is needed for his axiom of infinity. And he points out that Wittgenstein's notion of tautology cannot be applied to all the statements necessary for the deduction of classical mathematics.

To turn statements like 'Every vixen is a female fox' into logical truths you must be able to substitute synonyms for synonyms (in this case, 'female fox' for 'vixen'). But when are two terms synonymous? Like Professor Quine, White holds that the notion of analyticity cannot be clarified by the notion of synonymy, because the latter is as obscure as the former. You could say that two terms are synonymous if you could say that they convey or express an identical *meaning*; but White allows no appeal to such a metaphysical entity.

White now considers a criterion of analyticity which appears to dispense with the dubious notion of synonymy, namely that a statement is analytic if and only if its denial is self-contradictory. This works in the unproblematical case of logical truths. But how, in order to prove the analyticity of 'All men are rational animals', can we prove the self-contradictoriness of 'Some men are not rational animals'? Only by declaring that 'man' and 'rational animal' convey an identical *meaning*, or that these two words are *synonymous*. But both these moves have already been disallowed.

Can analytic truths be singled out by appealing to the linguistic fact that no one who understands the language would call a non-instance of the subject-term an instance of the predicate-term, or vice versa? For example, can the analyticity of 'All men are rational animals' be established by the fact that no one would call something a man which was not a rational animal? White argues that this procedure breaks down because, if we take terms which are regarded as having the same denotation or extension but a *different* connotation — 'man' and 'featherless biped', for instance — then it is also a fact that no one would call something a man which was not a featherless biped although 'All men are featherless bipeds' is, *ex hypothesi*, not analytic.

The conclusion towards which these animadversions are leading is that, logical truths apart, so-called 'analytic' statements cannot be *described* as objectively and inherently analytic; they can only be *made* analytic by stipulation. But before he arrives at this conclusion, White first rejects the quasi-conventionalist or half-way view

that 'All men are rational animals' is analytic because ordinary language behaves *as if* 'man' and 'rational animal' had been made synonymous by a linguistic convention. Like Bentham, he refuses to admit an appeal to an imaginary social contract. Either there was no stipulation, in which case 'man' and 'rational animal' have *not* been made synonymous, or else someone did stipulate that 'All men are rational animals' is analytic in *his* language. Stipulative procedures are artificial, and you cannot invoke a ghostly version of such procedures to explain alleged features of natural languages.

The artificial creation of synonymy is a performance which, like other wilful performances, requires to be justified. (Again, philosophy is a quasi-ethical activity of justifying linguistic legislation.) White doubts whether analytic-making activities serve any useful social purpose; and he emphatically denies that there are any factual considerations which ever *require* sentences to be rendered analytic. He uses an interesting analogy to drive home this point. To a lawyer a man is legally guilty only if he has been pronounced 'guilty' by a jury. But the jury will not *pronounce* him 'guilty' unless they are convinced by factual considerations that he *is* guilty. But, White argues, philosophers have turned to the idea that a sentence is analytic only if it has been pronounced 'analytic' precisely because they have *failed* to find factual considerations which are sufficient to convince them that the sentence is as a matter of fact analytic. There are no 'pre-established synonymies' in natural languages which antedate, and justify, nominal definitions.

Ethics

So far, the rejection of meanings has led to the rejection of the idea of natural synonymy and so to the rejection of the view that 'All bachelors are unmarried', say, is inherently analytic. We must now consider how the rejection of meanings affects the distinction between factual statements and ethical utterances.

Once more, White begins by considering a platonic ontology — in this case, the ontology of Dr. G. E. Moore's *Principia Ethica*. Here, the distinction between descriptions and evaluations was grounded (as the distinction between contingent and necessary statements had been grounded) on the existence of meanings which can be grasped or intuited, and which are neither physical nor psychical. Here, the process of evaluating something was said to involve an understanding of non-natural predicates, just as the process of describing something had been said to involve an understanding of natural predicates.

White rightly says that it is important 'to realize the degree to which Moore opposed what might be called a linguistic view of

philosophical analysis'. He was not concerned with how people *use* the word 'good'. Maybe they use it and 'pleasurable' in an interchangeable way; if so, this may be because 'good' and 'pleasurable' have the same extension, denote the same things. But none of this would have disturbed Moore, because he was only concerned to show that what is *meant* by 'good' cannot be what is *meant* by 'pleasurable.'

Moore's radical distinction between natural and non-natural predicates also involves the idea of analyticity. For one of Moore's ways of asserting the autonomy of ethical predicates was to say that no assertion of the form 'Everything which is good is *p*' (where '*p*' stands for some natural property or descriptive predicate) is analytic. This contention was based by Moore on the assertion that 'good' is indefinable; and this in turn was supposed to be a reflection of the ontological fact that goodness is a *simple* quality.

No prize is offered for anticipating White's attitude to a postulated entity which is not only occult and mysterious but simple and unanalysable as well! 'Good', White says, *can* be defined — in any way you please. You *may* declare, 'I take "good" as undefined', but that is a performative utterance and not a statement about the simplicity of what is allegedly connoted by 'good'. Nor, says White, can Moore gain any support for his contention that 'good' is indefinable from Bishop Butler's truism which Moore took for his motto; for the fact that a brother is what it is and not another thing does not stop 'brother' being definable as 'male sibling'. Moreover, just as it is perfectly possible to ask whether a male sibling *is* a brother, so, White contends, it would be perfectly possible to ask whether pleasure *is* good if 'good' had been defined as 'pleasurable'. Thus Moore's contention that if 'good' were defined naturalistically it would no longer be possible to ask whether the *definiens* is itself good collapses. But I imagine that Moore would reply that a correct affirmative answer to this question would be tautological and uninformative, whereas what is wanted by someone who is hesitating over, say, utilitarianism and who asks whether happiness really is the sole good, is an ethically *informative* reply; but a synthetic, informative reply, here, Moore would continue, is just what a naturalistic definition of 'good' renders impossible.

This last criticism apart, it seems to me that there is no defence against White's criticisms of Moore once the idea that words and sentences express extra-linguistic meanings and propositions has been repudiated.

For White, 'definable' and 'undefinable' are like 'desirable' and 'undesirable' in that they are prescriptive words which are easily

mistaken for descriptive words. Instead of saying that the simplicity of what is connoted by 'good' renders 'good' undefinable as a matter of fact, we should say that 'good' is undefinable in the sense that, although it *could* be defined, it would be inadvisable, or fruitless, or undesirable to define it. Thus the philosophy of ethics, like other branches of philosophy, is itself an ethical, or at any rate a quasi-ethical, activity.

We now turn, as we did in considering 'what is' and 'what must be', from platonism to positivism — in this case, from Moore's simple non-natural properties to the emotive theory of ethics. At first sight, everything seems different. Moore's objective, intuitable values have degenerated into subjective feelings; true-or-false ethical statements have become hortatory manifestations of emotion. But White holds, I think rightly, that the emotive theory is a descendant of Moore's theory and has inherited most of the latter's distinctive characteristics. The distinction between natural and non-natural predicates is here in a new guise as the distinction between cognitive and emotive meaning; and positivists, White says (though he should have excepted Schlick), are as hostile as Moore to ethical naturalism.

Just as he found the earlier logical positivists' criterion of *meaningfulness* unconvincing, so White now finds the moral positivists' *theory* of meaning unconvincing — the psychological theory according to which the meaning of a word is its disposition to arouse certain responses in the reader or hearer, the meaning being cognitive if the responses involve beliefs, and emotive if the responses involve attitudes.

White is as suspicious of these, to him, mysterious and occult dispositions or 'virtues' of words as he is of platonic 'meanings'. His quasi-ethical conception of philosophy has the effect of turning the tables on positivists who declare that ethical utterances are 'cognitively meaningless'; for such a declaration is, according to him, performative rather than descriptive, and therefore stands in need of the kind of justification which it vetoes.

White's own analysis of 'I approve of this' is that 'it is very close to being performatory in just the sense in which "I promise" is performatory and in the sense in which the logical legislator's "I define" is performatory'. I agree that 'I approve of this' is sometimes said in a performative way by people in authority. But the contention that the normal moral use of 'I approve' is performative seems to me to run into big difficulties. In the first place, as Austin has pointed out, it is no use making performative utterances if you are not in a proper position to do so. The Home Secretary may say 'I approve of the

appointment of X as the governor of this prison' in an authoritative and performative way; but it would be absurd for a junior civil servant who had no say in this appointment to stamp it with *his* approval. Yet he might very well approve of the appointment in the sense that he privately believes that the best candidate had been chosen. Secondly, this interpretation transforms a person's approving of something into an *occurrence*, or at best a series of occurrences. (Indeed, White is committed to this by his suspicion of *anything* invisible and dispositional.) But I just do not think that my approval of, say, candour in argument can be analysed without remainder into a series of dateable endorsements. If I were asked when I last stamped candour with my approval I should not know what to say. Thirdly, it seems to me that this interpretation involves a vicious infinite regress. For White insists that acts of approving, like other performances, require justification. If our Home Secretary were asked why he had 'approved' (in the performative sense) of the appointment of X he might reply that he had done so because he approved (in a non-performative sense) of the appointment of X (just as, to use White's analogy, a jury's justification for *pronouncing* a man 'guilty' will be that they *believed* him guilty). But if ordinary moral approving is also performative, our Home Secretary *cannot* offer this justification. And since approval seems to be necessarily involved in any full moral justification, there seems no escape from the conclusion that White both insists that performances should be justified and (since a performance cannot be justified by a further performance) makes it impossible to justify them.

However, White might retort that this criticism presupposes that a moral justification must contain ethical premises, a presupposition he seems to reject when he discusses the question, Is there an unbridgeable gulf between descriptive statements (e.g. 'I promised to do so-and-so and there is nothing to stop me doing it') and the prescriptive statements they are held to justify (e.g. 'I ought to do so-and-so'). In the course of a discussion of Mr. Stuart Hampshire's paper, 'Fallacies in Moral Philosophy', White points out that the feeling that there is a gulf can be intensified by talking about descriptive *statements* on the one hand, and moral *judgments*, *decisions*, or even moral *acts* on the other, and diminished by talking uniformly about judging so-and-so to be the case and also judging that such-and-such ought to be done, or perhaps even *deciding* that so-and-so is the case and *deciding* that such-and-such ought to be done.

His next step is to point out, rightly, that ordinary language lends no support to the idea of a radical difference between descriptive and prescriptive language. I can say that it is true that, or a fact that, or

that I know that, such-and-such ought to be done. It may be added, however, that a recognition of this fact-like character of ordinary moral language has not stopped Professor Ayer, for example, from insisting in his *Philosophical Essays* (pp. 231-2) that so-called moral 'facts' are utterly different from empirical facts. There are two different things which the fact-like character of ordinary moral language might be held to imply. The extreme implication would be that to say that such-and-such ought to be done is to say the same *kind* of thing as to say that so-and-so is being done: moral facts simply exist and can be found out in the same way that other facts are found out. The other more modest implication would be that moral beliefs are like empirical beliefs only in this, that they too are right or wrong, true or false. I think that children, primitive peoples, military men and certain religious communities often believe in the first kind of, so to speak, brute moral fact; but most conscientious people, in the West at any rate, believe only in moral 'facts' in the second sense.

White does not try to make prescriptive statements derivable from descriptive statements by turning moral principles into peculiar rules of inference. He points out, in accordance with his general tendency to soft-pedal the differences between moral and factual arguing, that it is equally possible and unhelpful to turn scientific theories into 'rules of inference' which allow inferences from singular statements to singular statements. A justification or argument which begins with factual premises and ends with moral conclusions is admittedly incomplete, but no more incomplete than many common-sense explanation-sketches:

> Very often we are asked to explain why a is Q, and reply 'because a is P' — even where we cannot justify the generalization 'All P is Q'. We often feel that there is *some* connection between a's being P and a's being Q without being able to specify the general law or laws connecting them. . . .

The Upshot

Negatively, the effect of repudiating 'meanings' is to remove the ontological basis for any natural or objective divisions of sentences into meaningful and meaningless, analytic and synthetic, descriptive and prescriptive — we are advised by White to abandon the idea that logic, science and ethics are three autonomous disciplines, each with a peculiar mode of argument, and to regard them instead as parts of a sort of federal union — of a whole intellectual system whose job is to lead to 'successful prediction in science and smooth decision in

morals'. (I should have thought that something a bit more heroic than smooth decisions might occasionally be expected even from a career-politician.)

White's positive view might be described as an expanded 'Duhemism', an expanded version of the view that an experiment which misfires casts doubt, not on any single hypothesis, but on the whole body of theories and assumptions which gave rise to the falsified prediction (including the theories and assumptions employed in interpreting the behaviour of the experimental apparatus). White's big additions to this view are the inclusion, first of logic, mathematics and ethical principles, along with scientific and ontological theories, among the revisable theoretical components of the system, and secondly of feelings of moral horror, along with disappointing observations, among what might be called the obstinate base-line experiences with which the system as a whole has to be squared.

> Just as Duhem's view, when pushed to the extreme, makes it difficult to maintain a radical separation between the analytic and the synthetic . . . , so the view we have advocated will break down the remaining dualism between logic-*cum*-empirical science and ethics.

Since there are no ultimate differences of kind between logical, factual and ethical doctrines, and no ultimate differences of kind between sense-experiences and moral feelings, *any* of the former may be revised if the whole system is challenged by either of the latter.

This holistic view is not only entailed by the abandonment of platonic meanings but, according to White, encouraged by what he regards as the natural successor to the platonic theory, namely the Wittgensteinian theory of meaning. For if to understand an expression *E* is to know how to use *E*, if its meaning can only be explained by examining its use in various specimen situations, then a full account of its meaning 'would present the whole of the language-system or scientific theory or moral code in which the expression *E* figures as a part'.

The idea that everything in an intellectual system is revisable places a heavy onus on whatever it is with which the whole system has to accord. I have reported White as indicating that it is observational and moral experiences which occupy this privileged disciplinary status. But having accepted the need for what he variously calls 'temporary absolutes', 'terminal sentences', 'pinned-down sentences', or 'categorically acceptable statements', White later confesses that he can find no sure rule for distinguishing such relatively hard-and-fast, recalcitrant touchstones from the relatively revisable and

theoretical and accommodating components of the system. Just as his inability to find a definitive criterion for analyticity led him to abandon even a rough distinction between analytic and synthetic, so his inability to find a definitive criterion for basic statements leads him to what I consider the disastrous conclusion that *any* parts of the system may be 'pinned down' and the rest of the system adjusted thereto: 'the pinned-down is a random sample of the totality of our beliefs'. We are free to make alterations wherever we like, so long as we restore internal coherence.

The old and perhaps a little naive but still, I think, basically sound view was that speculative scientific theories should be tested against lower-order experimental generalizations, and ultimately against experimental observations. No doubt experimental generalizations and observations are impregnated by theory and interpretation and are, indeed, occasionally false; but in the event of a collision the initial presumption is in their favour. On this view, theories are like bold adventurers, and experimental generalizations are like humble policemen, a few of whom come unstuck from time to time.

Now White not only says that any part of the system may be 'pinned down'. He also says that whatever we do pin down is rendered immune to overthrow by being *made* 'true' by definition. But a statement which is true by definition is compatible with any synthetic statement, or synthetic system of statements, whatever. In other words, we can enrol anyone we like as a policeman, and whoever we do enrol we at once render impotent by depriving him of his right to prosecute. Synthetic statements, White implies though he does not say so, are not subject to any disciplinary control whatever.

Why should the terminal or basic statements be *made* true and thereby *vacuous*? Why should they not be synthetic and (much more often than not) true in the sense that they correspond with facts? The answer is that White's anti-platonism compels him to say that there are no 'facts' with which basic statements 'correspond'. He pours scorn on 'platonic "facts" out there' and condemns 'an invidious but unjustified hypostatization of "facts" to which scientific sentences correspond'. So far as I can make out, White replaces the idea of a statement corresponding or clashing with facts by the idea of a sentence cohering with or contradicting other sentences. Which sentences get accorded a privileged status must be a matter of choice since there is no extra-linguistic criterion to tell us which we ought to pin down. The rejection of all 'occult' extra-linguistic entities leads, in the end, to a sort of linguistic solipsism: I am for ever imprisoned within my language-system. There is no exit from words to meanings, propositions, or facts. The meaning of a sentence is nothing but

its use and can be explained by nothing but a parade of further sentences, and its validity depends on nothing but its relations with other sentences. The world has disappeared behind an organic system of mutually self-accommodating sentences, outside which there are no facts, no fixed points, to which anything within this system has to correspond.

A curious transformation overcomes the concluding pages of this book. At first it seems as if there is to be a general tightening-up of discipline: more kinds of statements are to be exposed to a wider range of corrective experiences. To take an extreme example, moral repugnance may lead to the overthrow of a logical principle. But just this fact that *any* ingredient *may* be revised means that *no* ingredient *need* be revised. The number of potential scapegoats has been made so gratuitously large that no member of the system need feel culpable. If an ethical-factual system suggests a morally repugnant course of action, the moral principle can point an accusing finger at the scientific theory, and the scientific theory can put the blame on logic. What happens if logic is found guilty White does not tell us and I cannot imagine. Finally, it transpires that there is nothing outside the system with which it can correspond.

'The ship of science and morals', White says, 'is not moored by any easily specified set of beliefs or sentences.' And its anchors never get a grip on the bottom.

Do Intangible Entities Exist?

All these unfortunate conclusions are held by White — I think rightly — to follow from the rejection of intangible entities, especially meanings. Samson, it will be remembered, was inspired to destroy the temple because he had been blinded. White, too, has been blinded to the existence of anything intangible — meanings, propositions, dispositions, facts. And the result is that the distinctions between analytic and synthetic, between descriptive and prescriptive, between the act of stamping with approval and approving in a dispositional way — all these distinctions together with the idea that true statements correspond with facts have crashed about his ears.

His reasons for rejecting meanings were that they are (*a*) occult entities, and (*b*) postulated by philosophers in an *ad hoc* and non-explanatory way. He regards both occultness and *ad hoc*ness as sins of which physics is free. Yet powerful explanatory scientific theories postulate various entities far more strange and occult than anything brought to the notice of the Society for Psychical Research — gravitational and magnetic fields, positrons, unconscious wishes, marginal utilities, and so forth. The occult holds no terrors for

scientists though, of course, they would entirely agree that occult entities should not be postulated *ad hoc* as mere duplicates of observed phenomena.

This brings us to White's second complaint. Platonic 'meanings' introduced to 'explain' why we understand words and recognize the *a priori* character of necessary truths have, he frequently says, a status analogous to that of the 'dormitive virtue' introduced to 'explain' why opium induces sleep. If 'dormitive virtue' means 'sleep-inducing tendency', then 'opium tends to induce sleep because of its dormitive virtue' is certainly a purely *ad hoc* explanation of the form '*P* because *P*'. Since this *explicans* merely repeats the *explicandum*, the *explicans* can be eliminated without making any difference to our inventory of the word. Would the elimination of meanings similarly make no difference? White's book dramatically shows that it makes an enormous difference. The consequences he draws from the rejection of platonic meanings add up to an overwhelming refutation of his claim that meanings should be rejected because they are useless and mysterious duplicates of familiar processes. White will reply that the elimination of meanings has implications only for philosophical opinions and not for the familiar mental activities of ordinary people. But suppose that a layman uncorrupted by philosophy is confronted by these three sentences:

(1) All bachelors are unmarried.
(2) Some bachelors have mistresses.
(3) No bachelors ought to have mistresses.

I am prepared to lay a fair-sized bet that after some reflection and questioning he will realize that a different *kind* of justification is appropriate to each of these sentences. For instance, I am confident that he would admit that the findings of a survey of Bloomsbury might justify (2) but not (1) or (3), and that an understanding of the meanings of the words used in these sentences would by itself establish (1) but not (2) or (3). Now on White's own showing, the analytic status of (1), the synthetic status of (2), and the prescriptive status of (3) *could* be accounted for *if* meanings were postulated; whereas, if meanings are rejected, serious differences of kind between (1), (2) and (3) can no longer be accounted for. The 'platonic' theory of meaning, far from being *ad hoc*, explains what White's rejection of meanings renders incomprehensible, namely the layman's confident discrimination between (1), (2) and (3).

And is the 'grasping' of intelligible but intangible meanings so strange and metaphorical as White alleges? I am confronted by the German word *Säugetier* and by a sample of the objects it denotes,

but something eludes me — its meaning. Seeing an elephant, a rhinoceros and a kangaroo, I first conjecture that it means 'large land animal'; but then I notice a mouse, and modify my guess to 'four-legged creature'. Then a whale draws attention to itself by blowing off, and I try 'living creature'. Then I am informed that sharks are *not* denoted by *Säugetier*. This foxes me. Then I suddenly believe I've got it. 'Mammal?' I inquire, and I learn that I am right. At 11.45 a.m. on Monday I grasped the meaning of *Säugetier*. This dateable occurrence gives me the dispositional ability to use the word *Säugetier* appropriately. But White (despite his suspicions of dispositions and his analysis of *approval* as a series of dateable occurrences) is committed by his rejection of meanings (and hence of any acts of grasping meanings) to the widely held view that to understand an expression is *nothing but* knowing how to use it. This view makes it hard to explain why children so often catch the meaning of a word *instanter*, for most know-how is acquired gradually. Is the sudden acquisition of the meaning of a word supposed to be like learning to ride a bicycle, a skill whose sudden acquisition is not preceded by any intellectual illumination?

In conclusion, I want to object to White's *mode* of attacking the notion of analyticity, for it seems to me that this mode of attack is illegitimate in that it is *bound* to succeed against any philosophical notion whatever. White himself is aware of this kind of objection but brushes it aside. I will quote his own words:

> He [Mr. Benson Mates] correctly observes that some philosophers will not accept the terms 'analytic', 'possible', 'necessary', 'self-contradictory', 'synonymous', 'meaning', and others unless at least one of this group of terms is defined exclusively by reference to terms outside of this group. And then he comments, 'It may easily happen that the conditions of adequacy for a definition are so strong that no adequate definition is possible . . .' This is a very discerning observation, but one is tempted to reply that it does more to discredit the whole circle in which the term 'analytic' travels than it does to condemn the narrowness of that term's critics.

I feel no such temptation. On the contrary, it seems to me that Mates has gone to the heart of the matter. Suppose we have a group of inter-related concepts which, taken as a whole, are considered *sui generis*. If all these concepts are to be rejected unless at least one of them can be satisfactorily defined, and if a definition is here regarded as satisfactory only if the defining concepts are *outside* this whole group, then these concepts cannot be 'satisfactorily' defined without

giving them a new and alien meaning.[3] Suppose I claim that the whole group of ethical terms is meaningful and, taken as a whole, *sui generis*. You refuse to allow this joint claim unless I can give at least one of these terms a precise definition. I set about defining, say, 'right' in terms of 'ought', 'duty', etc., but you complain that you find *all* the terms in this ethical group hopelessly obscure and that only a definition using terms which you do understand and which are therefore right outside this group will satisfy you. Now you've got me, whichever way I wriggle. If I fail to comply you triumphantly reject my claim that ethical terms are meaningful. If I do comply I destroy my claim that ethical terms are *sui generis*. I must either surrender the meaningfulness of 'right', 'ought', etc., or commit the naturalistic fallacy. Similarly, White seems to me to require the defenders of the notion of analyticity either to admit that it is hopelessly obscure or to so define it that it means something they never intended it to mean.

To conclude that I regard this as an extremely valuable and educative book, a real eye-opener, may sound like a mild redress for the hard things I have said. But I mean it. For some time now philosophers have been trying to break down the 'analytic/synthetic' dichotomy. At first I regarded such attempts with much sympathy. Later I came round to the view that an investigation of the divisions *within* the heterogeneous class of synthetic statements is likely to be much more fruitful, but I still saw no danger in attacking analyticity. Again, it has been fashionable for some time now to say that the meaning of an expression is nothing but its use, and to regard talk about the 'propositions' expressed by sentences as dreadfully metaphysical. I regarded this unsympathetically, but I saw no particular danger in it. Now Professor White, with real deductive power and by boldly accepting whatever his premises implied, has shown us what these related tendencies lead to.

[3] Messrs. H. P. Grice and P. F. Strawson have made a rather similar objection to Quine's way of disposing of analyticity. They point out that Quine first makes the inference 'from "We have not made satisfactory sense (provided a satisfactory explanation) of *x*" to "*x* does not make sense" '. They then point out that, given Quine's exacting specifications for a 'satisfactory' explanation, 'it is perhaps dubious whether *any* such explanations can *ever* be given' (*The Philosophical Review*, April 1956, p. 148).

ON THE OPERATIONAL FOUNDATIONS OF LOGIC AND MATHEMATICS

By Gert H. Müller

In a series of publications, and particularly in his book *Einführung in die operative Logik und Mathematik* (Introduction to Operational Logic and Mathematics, Springer Verlag, 1955), Paul Lorenzen has proposed a new approach to foundational problems which — as can be seen from the title of the book — is related to recent attempts in philosophy to minimize the number of material ('inhaltliche') assumptions. Generally this trend towards elementarization in the philosophy of science is to be welcomed, and the objections raised in this paper are merely directed against attempts to carry the method too far.

Let us first recall the situation which arose during metamathematical investigations of the foundational problems. In some branches of mathematics, systems were laid down axiomatically and to a certain extent justified by their material interpretation. Since, however, such an interpretation was felt to be insufficient in assuring consistency, the need arose for proofs of formal consistency in formalized systems, i.e. axiomatic systems with formalized logic. Here interpretation was ignored and only the formal properties of the expressions used or generated were considered. However, the carrying out of these metamathematical investigations required a kind of reasoning which soon proved to be that used in the theory of numbers — as a (materially) interpreted system. The foundations for this had then in fact to be sought in material considerations. For purposes of metamathematical investigations it soon proved insufficient to rely on 'intuitive' or recursive number theory alone, and increasingly stronger assumptions had to be made.

It is natural then to make the foundations of logic, and of the theory of numbers in general, the focal point of investigations, and to try to demonstrate their material consistency without first having recourse to a consistency proof which, in principle, does no more to establish a theory than is assumed in carrying it out.

Lorenzen has set himself the task of providing such a foundation first for logic, and then for mathematics. As regards logic, one might be tempted to found it on everyday usage or the intuitive handling of the concepts 'true' and 'false'; and as regards arithmetic, to rely on intuition and on some type of inductive thinking. This is, in fact, the path followed by Lorenzen. On closer analysis, however, it proves to

be a much more complicated path and one involving many more assumptions than might have been expected.

Lorenzen first develops (unfortunately in too sketchy a manner) a theory of calculi. Calculi are directions for the generation of figures (expressions); for instance, we may call O, + 'atoms' of the calculus C, $\rightarrow$ O + a 'beginning', i.e. a rule for 'making' this figure (i.e. for regarding it as derived) and $\mathfrak{a} \rightarrow \mathfrak{a}$ O +[1] a rule for deriving the expression to the right of the arrow if the expression to the left of the arrow resulting from substitution has already been derived. Thus the calculus C generates the expressions O+, O+O+, O+O+O+, etc. We may symbolize the set of expressions derivable in C by $|\overset{\wedge}{\mathsf{C}}$. By the introduction of object-variables, a more complicated form of calculus is then constructed out of a calculus set up beforehand. For instance, let N be the calculus with the atom |, the beginning $\rightarrow$ |, and the rule $\mathfrak{a} \rightarrow \mathfrak{a}$|, so that $|\overset{\wedge}{\mathsf{N}}$ comprises the figures |, ||, |||, etc., then Lorenzen considers a calculus D with the atoms |, =, the beginning $\rightarrow$ | = |, and the rule $\mathfrak{x} = \mathfrak{y} \rightarrow \mathfrak{x}| = \mathfrak{y}|$, which generates the expressions | = |, || = ||, ||| = |||, etc. Here $\mathfrak{x}$, $\mathfrak{y}$ are object-variables, whose range is $|\overset{\wedge}{\mathsf{N}}$, in the calculus D. Lorenzen considers even more complicated types of calculi, where the derivation of an expression requires a certain set of other expressions to have been derived beforehand; but we shall not dwell on this point.

Lorenzen then examines the concept of derivability. An expression A in a calculus C is derivable in C ($\vdash_{\mathsf{C}}$A), if by successive applications of the rules it can be obtained from the beginnings. Whether a particular procedure of this kind is a correct derivation of A in C is decidable, at least in the case of simpler calculi. While this decision can be made by a simple comparison of symbols, there is generally no mechanical way of deciding whether A is derivable in C.

The judgment $\vdash_{\mathsf{C}}$A is proved by producing a derivation of A in C. $\nvdash_{\mathsf{C}}$A (A 'non-derivable' in C) is proved by a demonstration that every expression in $|\overset{\wedge}{\mathsf{C}}$ is semiotically different from A. Semiotic comparisons having been necessary for determining whether anything is a derivation, one might have expected Lorenzen to introduce the relations of semiotic equality and inequality with reference to intuition. This would have been quite correct methodologically since, already in introducing the concepts of the calculus and of derivability by mere examples, reference was made to intuition and this is subsequently used for justifying certain principles. It is presumably due to Lorenzen's too rigid adherence to a formal or 'operational' point of view that he sets up a special 'equality-calculus' — E_{C} — for a given calculus C, used in generating the

[1] Gothic letters are used here since these variables belong to the metalanguage.

expressions $\mathfrak{A} = \mathfrak{B}$ or $\mathfrak{A} \neq \mathfrak{B}$. Here $\mathfrak{A}$, $\mathfrak{B}$ refer to any given combination of signs in C. In order to recognize E_C as an equality-calculus, i.e. in order to evaluate the figures $\mathfrak{A} = \mathfrak{B}$, etc., as 'correct' semiotic equalities, judgments of the type $\vdash_{E_C} A$ (e.g. $\vdash_{E_C} \mathfrak{A} = \mathfrak{B} \wedge \mathfrak{A} \neq \mathfrak{B}$) must first be proved, for which — apart from the extension due to $\wedge$ to which we shall return later — a new calculus $E^E{}_C$ is required, and so *ad infinitum*. On the whole, Lorenzen's treatment of equality is not the happiest chapter in his book.

A rule R is called 'admissible' in a calculus C if its addition to the rules of C does not extend $|\frac{A}{C}$. An admissibility judgment involves a judgment about the totality of derivable expressions in a calculus, and, in order to establish this judgment, certain 'protological' principles are propounded and examples are given to make them evident.

First the principle of induction is considered, but it is not formulated clearly enough, especially with regard to its applications. Its substance may be stated in the following way: if a 'property' is 'derivable' for all beginnings of a calculus C, and if its hereditary nature for all the rules of C can be proved 'admissible', then all figures derivable in C have this property. It is left open to which calculus C the terms 'derivable' and 'admissible' refer, and what is to be understood by 'property'. This principle is supposed to cover the cases in which inductive arguments are employed for the figures in $|\frac{A}{N}$ (later to be interpreted as numbers), but also metamathematical inductive arguments for the derivability of expressions in calculi, and furthermore, inductive arguments for properties such as 'stable in van Dantzig's sense', 'semiotically different from a figure', or 'definite' — a concept which we will discuss later.

Another principle is the principle of non-derivability. It states that if an expression A is non-derivable in a calculus C, then the rule $A \rightarrow B$ (where anything whatsoever can be substituted for B) is admissible in C, which is evident since A can obviously not be used as an 'assumption'.

We shall not embark on the discussion of Lorenzen's three further principles for making admissibility judgments.

In considering this theory which Lorenzen calls *protologic* in which the possibilities of proving judgments of the kind '$\vdash_C A$', '$\nvdash_C A$', 'admissible $_CR$', or 'non-admissible $_CR$' are examined, we are naturally reminded of similar investigations by G. Gentzen and his followers, or by H. B. Curry. However, Lorenzen's development of his theory and his introduction of admissibility judgments are undoubtedly original. Unfortunately the development of his theory of calculi and

of protologic is impaired by a confusing anticipation of the later contents of the book, by an inadequate introduction of the different types of variables, and by too sketchy and somewhat over-generalized formulations of the principles.

Although, in practice, Lorenzen's concept of definiteness hardly enters into his actual development of logic and mathematics, he nevertheless seems to regard it as important. However, his explanations of the concept are rather equivocal.

We call a judgment $\vdash_c A$ 'definite', or more precisely, 'proof-definite', because we know how to prove it; and we call the judgment $\nvdash_c A$ 'refutation-definite' because we know how to refute it, viz. by stating a derivation of A in C. Similarly we realize that also judgments about the admissibility or non-admissibility of a rule in a calculus are definite in the sense that it is laid down how they are to be proved or refuted, i.e. under what conditions we may make these judgments. Thus this concept of definiteness is only applicable to judgments and not to figures of a calculus such as Fermat's Last Theorem — an expression in a calculus of the theory of numbers. Also questions such as whether the proposition 'x is a set' is definite or not, are meaningless. Furthermore, the rather vague inductive definition:

'(1) All propositions which can be decided by schematic operations are to be considered as definite;

'(2) if a definite concept of proof or refutation is stipulated for a proposition, then the proposition itself is to be considered as proof-definite or refutation-definite',

is hardly necessary in this connection. Finally, judgments of derivability or refutability are obviously definite even in strongly impredicative (formalized) set theories, so that there can be no question of excluding impredicativity as a result of demands for definiteness concerning judgments.

However, it is not clear how the concept of definiteness is to be introduced so as to be applicable to judgments, sentences (i.e. expressions in a calculus), and interpretations ('x is a set, a number or a calculus', etc.).

When applied to judgments, the demand for definiteness does not immediately lead to limitations, since the task of protologic is precisely to ensure that the judgments in question are definite. On the other hand, the demand for definiteness would exclude, or at least limit, principles such as Carnap's Rule or the related principle of 'infinite induction', which Lorenzen, in particular, has used in some of his papers. These principles allow the derivation of a universal proposition $\Lambda_{\mathfrak{x}} A(\mathfrak{x})$, if all statements $A(\mathfrak{x})$ are derived for every $\mathfrak{x}$.

In an infinite domain of individuals this requires the derivation of an infinite number of premises. Whenever such principles are not limited by metamathematical requirements, e.g. that these infinite sets of derivations be inductive in a strictly specified sense, the judgment $\vdash \Lambda_{\mathfrak{x}} A(\mathfrak{x})$ can generally not be considered as definite, even when the judgment $\vdash A(\mathfrak{x})$ is definite for every $\mathfrak{x}$.

Lorenzen then proceeds to develop logic and mathematics on the basis of protologic.

As the basic procedure in the foundation of logic he proposes that a series of formal calculi be set up which successively generate the 'correct' figures of pure consequential logic, positive propositional logic, positive quantification logic, and intuitionist (quantification) logic. The 'correctness' of the figures (propositions) generated by these calculi is to be demonstrated in part by interpreting these propositions as being judgments of derivability and admissibility, and in part by considering them as being uninterpreted, or as having uninterpreted parts that cannot be considered 'correct' in any sense.

Let us illustrate this procedure by two examples, one taken from consequential logic, the other from positive proportional logic. For setting up the former, we heuristically lay down some rules such as that of the transitivity of the arrow-operation:

(T) $\qquad A_1 \rightarrow A_2, A_2 \rightarrow A_3 \dot{\rightarrow} A_1 \rightarrow A_3$

By a method of elimination for any given derivation in which the rule $A_1 \rightarrow A_3$ was used, this rule may then be eliminated by applying $A_1 \rightarrow A_2$ and $A_2 \rightarrow A_3$. This leads to the conclusion that (T) is a 'generally admissible' meta-rule for every calculus. (Note that a tacit change in the meaning of the term 'rule' has here been introduced, for meta-'rules' in this sense do not generate uninterpreted expressions, rather $A_1 \rightarrow A_3$ is meant to be interpreted as a rule, for otherwise (T) would not be generally admissible but would simply be an ordinary rule for generating expressions of a specified type. This is only possible in specified calculi with specified atoms.) Subsequent meta-rules are similarly investigated, with a view to setting up consequential logic, i.e. a procedure for obtaining all such meta-rules. This is done by setting up a calculus CL, but this requires first the setting up of a calculus — we call it S_{CL} — which generates the syntactically required figures of consequential logic. Its atoms would be a series of variables $A_1, A_2 \ldots$, and also the signs '$\rightarrow$', ',' and brackets. Then the consequential logic calculus CL with object-variables ranging over $\mid\!\!\overset{\wedge}{-} S_{CL}$ and with a new atom, e.g. 'ua' (universally admissible) for generating ua[. . . . . .] could be set up. Now in order to justify the calculus CL an inductive proof, for example, would have to be carried out in let us say the following

way: 'The beginnings of CL can be regarded as universally admissible with such-and-such an interpretation, and the rules of CL allow this property to be inherited.' Here, and even more so when setting up later calculi, Lorenzen has neglected to carry out all these steps and to take into account the axiomatic character of his procedure, revealed in setting up calculi. In any case, we must emphasize that such a proof seems quite possible for CL, if arrow-expressions are interpreted as admissibility judgments.

For positive propositional logic, the formal calculus PL is set up by adding the atoms $\wedge$ and $\vee$ to a calculus C, and by adding rules such as $\mathfrak{a} > \mathfrak{b} \rightarrow \mathfrak{a} \wedge \mathfrak{b}$, $\mathfrak{a} \rightarrow \mathfrak{a} \vee \mathfrak{b}$, $\mathfrak{b} \rightarrow \mathfrak{a} \vee \mathfrak{b}$, but without setting up a parallel calculus S_{PL} for the generation of the required syntactical figures. In that case there can be no question of a justification since $\wedge$ and $\vee$ are uninterpreted signs. There is no reason why quite different rules could not have been chosen for these signs, and the calculus still be called positive logic of propositions. It must also be realized that by introducing $\wedge$ and $\vee$ into the calculus PL, the sign $\rightarrow$ has been given a new meaning so that a justification (foundation) of the *whole* calculus would be required. A general flaw in Lorenzen's exposition is his ignoring of the fact that with the introduction of each new calculus the entire new system of calculi and not only the particular additions must be justified. Moreover, if PL is to be brought into a specific association with another calculus such as enables it to be used as logic, then a proof of justification becomes all the more essential. True, in this case expressions are produced operationally, and this is quite a permissible procedure, but no foundations are given.

Then again additions due to the introduction of the universal quantifier and of negation are justified, as in the case of consequential logic, by interpretations based on admissibility judgments. On the other hand the addition of the existential quantifier $\vee\mathfrak{x}\ A(\mathfrak{x})$ is, as with the addition of $\wedge$ and $\vee$, purely formal and without interpretation or justification, with the result that the intuitionist 'logic' so obtained has a heterogeneous character and its application is far from clear.

The introduction of the universal quantifier is inadequate and contains a fallacious proof (p. 53). The idea is probably to justify the step from $A_0 \rightarrow A(\mathfrak{x})$—($\mathfrak{x}$ not free in A_0)—to $A_0 \rightarrow \wedge\mathfrak{x}\ A(\mathfrak{x})$ by considering $A_0 \rightarrow A(\mathfrak{x})$ as a set of rules, covering all such rules for each individual object. The step to $A_0 \rightarrow \wedge\mathfrak{x}\ A(\mathfrak{x})$ is made once this set of rules has produced the figures $A(\mathfrak{x})$ for each individual object $\mathfrak{x}$. A proof of the admissibility of $A_0 \rightarrow \wedge\mathfrak{x}\ A(\mathfrak{x})$ has thus to be made by means of a proof of the admissibility of a set of rules, and whether

such a judgment is definite depends, of course, on whether this set of rules is described in a finite way. Only for such sets of rules can the universal quantifier perhaps be considered as explained. (Lorenzen's exposition on this point is so obscure that every attempt at clarification is bound to be frustrated.)

A particularly impossible situation arises when attempts are made to justify the use of the negation. For a figure (sentence) A, the negation can only be formed in calculi in which non-derivable sentences are known. If Λ be such a non-derivable sentence, then the sentence $\neg A$ is interpreted as the rule $A \rightarrow \Lambda$. As long as one substitutes a 'false' sentence for Λ in a calculus (false in the sense of the intuitionists who regard this concept as intuitively clear), no difficulties arise; rather is it a case of the usual *ex falso quod libet*. However, an operationalist ought to explain the existence and the meaning of 'false', which seems impossible without recourse to interpretation. Now if instead of Λ one uses any sentence non-derivable in a calculus (for instance, Peirce's Law $[(A \rightarrow B) \rightarrow A] \rightarrow A$ in the calculus CL after introducing the negation into it) one is led to contradictions or more exactly to the derivability of Λ itself.

As in all such attempts to create a foundation of logic, the principal difficulty is not so much the setting up of calculi (generation-procedures) for the formation of the required figures, but the demonstration of certain formal properties of these calculi, and particularly the justification of the choice of such calculi. Especially in giving reasons why a calculus should function as a *logic* calculus in combination with mathematical calculi, an interpretation cannot be avoided; and this introduces an element which is different from and alien to operationalism.

Furthermore, Lorenzen allows an extension of intuitionist logic by admitting the *tertium non datur*, an extension, however, which he expressly designates as fictitious whenever it is applied. A certain justification for the use of the *tertium non datur* is seen in the method of interpreting classical logic and arithmetic in the systems of intuitionistic logic and arithmetic, a method which in essence goes back to Gödel.

It must be added that the introduction of the ι-terms (descriptions) $\iota\mathfrak{x}\, A(\mathfrak{x})$ and of the class-terms $\hat{\mathfrak{x}}\, A(\mathfrak{x})$ contains a methodological error. In the case of ι-terms, quite apart from inadequate formal treatment, the assumption that there exists between an object (e.g. the number 2) and one of its descriptions (e.g. the even prime) a semiotic equality, leads to difficulties. In the introduction of class-terms, the author uses class equivalence. Admittedly the equivalence $A(\mathfrak{x}) \leftrightarrow B(\mathfrak{x})$ allows a classification of the sentence-forms $A(\mathfrak{x})$

according to logical equivalence, and in mathematical terminology one can regard $\mathsf{A}(\mathfrak{x})$ as the representative of its equivalence class, but obviously *not* as the class of all $\mathfrak{x}$ for which $\mathsf{A}(\mathfrak{x})$ holds. For the latter, i.e. $\hat{\mathfrak{x}}\ \mathsf{A}(\mathfrak{x})$ is a term in which $\mathfrak{x}$ is an *apparent* variable, while in the sentence-form $\mathsf{A}(\mathfrak{x})$ to which Lorenzen wishes to reduce the classes $\hat{\mathfrak{x}}\mathsf{A}(\mathfrak{x})$, $\mathfrak{x}$ is free for substitutions. Obviously this introduction vitiates Lorenzen's treatment of the predicate ϵ, of finite sets and of the type-hierarchy. In particular, his treatment of ι-terms and class-terms shows that the operational methodology of generating and that of reducing to judgments of derivability, simply misses the point.

As regards the further development of mathematics — apart from the theory of numbers to which we shall return later — we here add the following: Lorenzen's aim is the development of *predicative* analysis, i.e. the type of analysis which can be carried out within the framework of the ramified theory of types proposed by Russell (without a reducibility axiom). By 'predicative' is meant that the introduction of a particular class occurs without reference to a totality of classes of which the class to be introduced is a member, or more briefly that the introduction of classes is in accordance with the vicious circle principle.

In following this programme difficulties arose due to the limitation to finite types particularly also in Weyl's approaches. These Lorenzen overcomes by admitting a type-hierarchy topped by type numbers characterized by transfinite constructible ordinal numbers, a procedure which, by the way, was also suggested by Hao Wang in an important contribution to the *Journal of Symbolic Logic*, 1954.

The development of predicative analysis on the basis of its mathematical content — i.e. without considering Lorenzen's way of foundation and formal treatment — gives a number of valuable results, and these are often presented in a simplified manner compared with the respective publications of the author.

Since Lorenzen considers predicative analysis to be the maximum of what is permissible or certain in mathematics, why then should he not delimit the permissible with the relatively simple and easily introduced concept of predicativity, rather than with the concept of definiteness which is much less clear?

The operational development of the theory of numbers suffers as much as that of logic from the fact that the mere making of expressions does not permit them to be interpreted as propositions.

The point of departure is the previously mentioned calculus N, which generates the expressions $|$, $||$, $|||$, etc. To this are added the calculi D, A, M, generating successively the expressions $| = |$,

$||=||, \ldots$ etc.; $|+|=||$, $|+||=|||, \ldots, ||+||=||||, \ldots$ etc.; $|\times|=|$, $|\times||=||\ldots., ||\times||=||||, \ldots$ etc., i.e. the 'correct' primitive propositions of equality and of the predicates of addition and multiplication. (The exact process of this development is obscured rather than clarified by the interpolation of a paragraph on 'strings of letters'.)

It must be emphasized that the successor operation has here been taken for granted (as being, in principle, the only function). This becomes clear from the fact that in the rule $\mathfrak{a} \rightarrow \mathfrak{a}|$ of the calculus N, $\mathfrak{a}|$ is correctly formed according to the definition of calculus, while in the rules of D, A, M, e.g. $\mathfrak{x} \rightarrow \mathfrak{y} \rightarrow \mathfrak{x}| = \mathfrak{y}|$, etc., $\mathfrak{x}|$ represents a formation unexplained by the general introduction of the concept of the calculus. $\mathfrak{x}$ is after all an object-variable and not a calculus-variable like $\mathfrak{a}$. This difference ought certainly to have been taken into account in attempting to found the theory of numbers.

Even more it must be emphasized that the calculi D, A, M start by generating uninterpreted expressions, and that a special justification is needed for considering these expressions (e.g. $| = |$) as significant assertions about the objects |, ||, etc. If intuition is invoked in this case, then the 'predicate' $=$ must have a material significance. This brings us back to Lorenzen's dilemma whenever he deals with equality. Additionally, the equality stipulated by the calculus D must be clearly distinguished from that stipulated by the calculus E_N (see above).

We must also realize that originally the calculi D, A, M have no inherent connections with one another. In order to make such connections, and to arrive at further statements protological principles are invoked. Hence the statements thus obtained ought really to be considered as metastatements, a distinction that Lorenzen (deliberately) omits. Protological principles are thereby endowed with the character of the usual principles of proof in an axiomatic system, and the role of axioms is taken over by the calculi as a whole. When protological principles are used as principles of proof, the addition of a new calculus to a previous one, e.g. of M to D and A, must necessarily represent a new demand on the 'numbers'. That these principles of proof play an essential part can be seen from the fact that only with their help can sentences (figures) like $\mathfrak{x}| = \mathfrak{y}| \rightarrow \mathfrak{x} = \mathfrak{y}$, $\mathfrak{x}| \neq |$ or $\mathfrak{x}+\mathfrak{y} = \mathfrak{z} \wedge \mathfrak{x}+\mathfrak{y} = \mathfrak{v} \rightarrow \mathfrak{z} = \mathfrak{v}$ be obtained.

Very little has remained of the operational point of view, which was supposed to reduce uninterpreted expressions to judgments of derivability or to judgments of admissibility. On the other hand the principle of replacing certain assumptions of normal axiomatic systems by setting up calculi represents another element of the

operational procedure which presumably led Lorenzen to consider mathematics a science of calculi. But this procedure by no means obviates the need for proving from the methodological point of view what is required in establishing the theory of numbers in an axiomatic way. Furthermore there are axiomatizations for number theory which instead of using axioms use schemata of axioms with (metamathematically used) numerals, and which are very similar to Lorenzen's procedure. The axiomatic existential element which Lorenzen thinks he has avoided lies essentially in the *combination* of the logic of quantification with the calculi D, A, M, *and* with the protological principles.

This interaction of three such heterogeneous possibilities of arriving at interpreted statements about numbers is obviously fraught with the danger of material contradictions. To guard against this danger, an appeal to intuition is clearly far less useful here than in the case of the usual introduction of number theory by interpreted axioms. Although this is at variance with his principal task, Lorenzen tries to adduce a kind of consistency proof which is bound to fail since it overlooks the combination to which we have referred. Furthermore the proof loses in methodological clarity through its simultaneous attempt to prove the above-mentioned result obtained by Gödel.

We shall not discuss further parts of the book, since a detailed discussion would take up too much space.

In the present analysis an attempt has been made to base it on Lorenzen's point of view; what basic criticisms have been advanced must be explained by the usual requirements made of logic and mathematics.

It will now seem obvious that any operational procedure in founding logic and mathematics must fail if it does not use material assumptions. For we cannot arrive at the characteristics of logic and mathematics if we proceed from a basis consisting of (1) the procedure to make figures according to different systems of rules (calculi) — just as we write according to a system of rules — and (2) an examination of the figures obtained by these systems of rules. We realize that for the process of derivation in any formalized theory, as for knitting a sweater, no more is required than what is assumed by (1) and (2). In fact, from the original figures new figures are systematically produced and the totality of the figures so obtained characterizes the theory (or the sweater). However, before we can single out logic or the theory of numbers from such 'theories', further assumptions are needed. There is no doubt that protological principles and the occasional implicit appeals to intuition represent material assumptions whose correctness cannot be judged formally,

and thus the operational framework is here already transcended. But as we saw in the case of the question, 'What is false?', and in the introduction of the class-terms $\hat{\xi}\, A(\xi)$, these assumptions are still insufficient.

It is important for a philosophy of science to define the position of logic and mathematics, and this position depends not only on their foundations but on the different possibilities of setting them up. From the failure of the attempt here discussed it emerges that even a comprehensive operational basis is inadequate for such a foundation. From the fact that most non-empirical sciences can be formalized, i.e. represented in the form of calculi (generally not without considerable loss in content), it simply does not follow that such calculi and operations within them will return us to these sciences.

2

ON THE STATUS OF SCIENCE AND OF METAPHYSICS

TWO RADIO TALKS

By Karl R. Popper

I. KANT AND THE LOGIC OF EXPERIENCE

In this talk[1] I do not propose to speak of ordinary everyday experience. I intend, rather, to use the word 'experience' in the sense in which we use it when we say that science is based on experience. Since, however, experience in science is, after all, no more than an extension of ordinary everyday experience, what I shall have to say will apply, by and large, to everyday experience also.

In order not to get lost in abstractions, I intend to discuss the logical status of a specific empirical science — Newtonian dynamics. I do not, however, presuppose any knowledge of physics on the part of my audience.

One of the things a philosopher may do, and one of those that may rank among his highest achievements, is to see a *riddle*, a *problem*, or a *paradox*, not previously seen by anyone else. This is an even greater achievement than solving the riddle. The philosopher who first sees and understands a new problem disturbs our laziness and complacency. He rouses us (as Kant said that Hume did for him) from our 'dogmatic slumber'. He opens out a new horizon before us.

The first philosopher clearly to apprehend *the riddle of natural science* was Kant. I do not know of any philosopher, either before or since, who has been so profoundly stirred by it.

When Kant talked of 'natural science' he almost invariably had Isaac Newton's celestial mechanics in mind. Kant himself made important contributions to Newtonian physics and he was one of the greatest cosmologists of all time. His two principal cosmological works are the *Natural History and Theory of the Heavens* (1755) and the *Metaphysical Foundations of Natural Science* (1786). Both books were (in Kant's own words) 'treated according to Newtonian Principles'.[2]

[1] The two talks are here reprinted almost exactly as they were given; only a few words have been changed, for stylistic reasons. The footnotes, of course, are additions.

[2] Also of great importance is the Latin *Physical Monadology* of 1756 in which Kant anticipated the main idea of Boscovic; but in his work of 1786 Kant repudiated the theory of matter propounded in his *Monadology*.

Like almost all of his contemporaries who were knowledgeable in this field, Kant believed in the *truth* of Newton's celestial mechanics. The almost universal belief that Newton's theory *must* be true was not only understandable but seemed to be well-founded. Never had there been a better theory, nor one more severely tested. Newton's theory not only accurately predicted the orbits of all the planets, including their deviations from Kepler's ellipses, but also the orbits of all their satellites. Moreover, its few simple principles supplied at the same time a celestial mechanics and a terrestrial mechanics.

Here was a universally valid system of the world that described the laws of cosmic motion in the simplest and clearest way possible — and with absolute accuracy. Its principles were as simple and precise as geometry itself — as Euclid's supreme achievement, that unsurpassed model of all science. Newton had indeed propounded a kind of cosmic geometry consisting of Euclid supplemented by a theory (which too could be represented geometrically) of the motion of mass-points under the influence of forces. It added, apart from the concept of time, only two essentially new concepts to Euclidian geometry: the concept of mass or of a material *mass-point*, and the even more important concept of a directed *force* (*vis* in Latin and *dynamis* in Greek from which the name 'dynamics' for Newton's theory is derived).

Here, then, was a science of the cosmos, of nature. And it was a science claimed to be based upon experience. It was a deductive science, exactly like geometry. Yet Newton himself asserted that he had wrested its fundamental principles from experience *by induction.* In other words, Newton asserted that the truth of his *theory* could be logically derived from the truth of certain *observation-statements.* Although he did not describe these observation-statements precisely, it is nevertheless clear that he must have been referring to Kepler's laws, the laws of the elliptic motions of the planets. And we can still find prominent physicists who maintain that Kepler's laws can be derived inductively from observation-statements, and that Newton's principles can in turn be derived, entirely or almost entirely, from Kepler's laws.

It was one of Kant's greatest achievements that, roused by Hume, he recognized the paradoxical character of this contention. Kant saw more clearly than anyone before or since how absurd it was to assume that Newton's theory could be derived from observations. Since this important insight of Kant's is falling into oblivion, partly because of his own contributions towards a solution of the problem he had discovered, I will now present and discuss it in detail.

The assertion that Newton's theory was derived from observation will here be criticized on three points:

First, the assertion is *intuitively not credible*, especially when we compare the character of the theory with the character of observation-statements.

Secondly, the assertion is *historically false*.

Thirdly, the assertion is *logically false*: it is a logically impossible assertion.

Let us examine the first point — that it is intuitively not credible that Newton's mechanics can, with the help of observations, be shown to be true.

To see this we merely have to remember how completely Newtonian theory differs from any observation-statement. In the first place, observations are *always* inexact, while the theory makes absolutely exact assertions. Moreover, it was one of the triumphs of Newtonian theory that it stood up to subsequent observations which, as regards precision, went far beyond what could be attained in Newton's own time. Now it is incredible that more precise statements, let alone the absolutely precise statements of the theory itself, could be logically derived from less exact or inexact ones. But even if we forget all about the question of precision, we should realize that an observation is always made under very special conditions, and that each observed situation is always a highly specific situation. The theory, on the other hand, claims to apply in all possible circumstances — not only to the planets Mars or Jupiter, or even to the satellites in the solar system, but to *all* planetary motion and to *all* solar systems. Indeed, its claims go far beyond all that. For example, the theory makes assertions about gravitational pressure inside the stars, assertions which even today have never been tested by observation. Moreover, observations are always *concrete*, while theory is *abstract*. We never observe, for example, mass points but rather extended planets. This may not be so very important perhaps; but what is of the utmost importance is the fact that we can never — I repeat, never — *observe* anything like Newtonian *forces*. Admittedly, since forces are so defined that they may be measured by measuring accelerations, we can, indeed, *measure* forces; and we may at times measure a force not by measuring an acceleration, but, for example, with the help of a spring balance. *Yet in all these measurements, without exception, we always presuppose the truth of Newtonian dynamics*. Without the prior assumption of a dynamical theory, it is simply impossible to measure forces. But forces, and changes of forces, are among the most important objects of which the theory treats. Thus we may assert that at least some of the objects of which the theory treats are

abstract and unobservable objects. For all these reasons it is intuitively not credible that the theory should be logically derivable from observations.

This result would not be affected even if it should be possible to reformulate Newton's theory in such a way that any reference to forces is avoided. Nor would it be affected by a dismissal of force as a mere fiction, or perhaps as a purely theoretical construction which only serves as a tool or instrument. For the thesis which we are questioning says that Newton's theory can be shown to be true by observation. And our objection was that we can only observe *concrete things*, while theory, and particularly Newtonian forces, are *abstract*. These difficulties are in no way mitigated if we make the theory even more abstract by eliminating the notion of force or by unmasking it as a mere auxiliary construction.

So much for my first point.

My second point was that it is historically false to believe that Newton's dynamics was derived from observation. Though this belief is widespread, it is nevertheless a belief in an historical myth — or, if you like, a bold distortion of history. To show this, I shall briefly refer to the part played by the three most important precursors of Newton in this field: Nicolaus Copernicus, Tycho Brahe, and Johannes Kepler.

Copernicus studied in Bologna under the Platonist Novara; and Copernicus's idea of placing in the centre of the universe the sun, rather than the earth, was not the result of new observations, but of a *new interpretation* of old and well-known facts in the light of semi-religious Platonic and Neo-Platonic ideas. The crucial idea can be traced back to the sixth book of Plato's *Republic* where we can read that the sun plays the same role in the realm of visible things as does the idea of the good in the realm of ideas. Now the idea of the good is the highest in the hierarchy of Platonic ideas. Accordingly, the sun, which endows visible things with their visibility, vitality, growth, and progress, is the highest in the hierarchy of the visible things in nature.

This passage in the *Republic* is of outstanding importance among the passages upon which Neo-Platonic philosophy — and particularly Christian Neo-Platonic philosophy — was based.

Now if the sun was to be given pride of place, if the sun merited a divine status in the hierarchy of visible things, then it was hardly possible for the sun to revolve about the earth. The only fitting place for so exalted a star was the centre of the universe. So the earth was bound to revolve about the sun.

This Platonic idea, then, forms the historical background of the

Copernican revolution. It does not start with observations, but with a religious or mythological idea. Such beautiful but wild ideas have often been put forward by great thinkers, and just as often by cranks. But Copernicus, for one, was no crank. He was highly critical towards his own mystical intuitions, which he rigorously examined in the light of astronomical observations, reinterpreted with the aid of the new idea. He rightly considered these observations to be extremely important. Yet looked at from a historical or genetical point of view, observations were not the source of his idea. The idea came first, and it was indispensable for the interpretation of the observations: they had to be interpreted in its light.

Johannes Kepler, Tycho Brahe's pupil and assistant, to whom his great teacher left his unpublished observations, was a Copernican. And, like Plato himself, Kepler was deeply influenced by the number-mysticism of the Pythagoreans. What he hoped to discover, what he searched for throughout his life, was the arithmetical law underlying the structure of the world, the law upon which rested the construction of the circles of Copernicus's solar system, and upon which, in particular, their relative distances from the sun were based. He never found what he sought. He did not find in Tycho's observations the hoped-for confirmation of his belief that Mars revolved about the sun in a perfectly circular orbit, with uniform velocity. On the contrary, he discovered in Tycho's observations a *refutation* of the circle hypothesis. Thus he discarded the circle hypothesis and tried the next best thing: the hypothesis of the ellipse. And he found that the observations could be made to agree with the new hypothesis — although only under the assumption, first far from welcome, that Mars did not travel with uniform velocity.

Historically, therefore, Kepler's laws were not the result of observations. What happened was that Kepler tried in vain to interpret Tycho's observations by means of his original circle hypothesis. The observations *refuted* this hypothesis, and so he tried the next best solution — the ellipse. The observations still did not prove that the hypothesis of an ellipse was correct, but they could now be *explained* by means of this hypothesis: they could be reconciled with it. Such, from an historical and genetical point of view, were the main antecedents of Newton's theory. Our story shows that as a matter of historical fact, the theory was not derived from observations.

Kant realized all this; and he also appreciated the fact that *even physical experiments* are not, genetically, prior to theories — no more than astronomical observations. They too simply represent crucial questions which man poses to nature with the help of theories — just as Kepler asked nature whether his circle hypothesis was true. Thus

Kant wrote in the preface to the 2nd edition of the *Critique of Pure Reason*: 'When Galileo let his balls run down an inclined plane with a gravity which he had chosen himself; when Torricelli caused the air to sustain a weight which he had calculated beforehand to be equal to that of a column of water of known height; . . . then a light dawned upon all natural philosophers. They learnt that our reason can understand only *what it creates according to its own design: that we must compel Nature to answer our questions*, rather than cling to Nature's apron strings and allow her to guide us. *For purely accidental observations, made without any plan having been thought out in advance, cannot be connected by a . . . law — which is what reason is searching for.*'[3]

This quotation from Kant shows how well he understood that we ourselves must confront nature with hypotheses and demand a reply to our questions; and that, lacking such hypotheses, we can only make haphazard observations which follow no plan and which can therefore never lead us to a natural law. In other words, Kant saw with perfect clarity that the history of science had refuted the Baconian myth that we must begin with observations in order to derive our theories from them. And Kant also realized very clearly that behind this historical fact lay a logical fact; that there were logical reasons why this kind of thing did not occur in the history of science: that it was logically impossible to derive theories from observations.

My third point — the contention that it is logically impossible to derive Newton's theory from observations — follows immediately from Hume's critique of the validity of inductive inferences, as pointed out by Kant. Hume's decisive point may be put as follows:

Take a class consisting of any number of true observation-statements and designate it by the letter K. Then statements in the class K will describe actual observations, i.e. *past* observations: thus we designate by the letter K any class whatsoever of true statements about observations actually made in the past. Since we have assumed that K consists only of true statements, all statements in the class K must also be consistent statements, and, furthermore, all statements belonging to the class K must be *compatible with one another*. Now take a further observation-statement which we shall designate by the letter B. We assume that B describes some *future, logically possible, observation*; for example, that B tells us that there will be an eclipse of the sun tomorrow. Since eclipses of the sun have already been observed, we can be certain that a statement B, asserting that there will be an eclipse of the sun tomorrow, is a statement which, on purely logical grounds, is *possible*; that is to say, our B is self-consistent. Now Hume shows the following: if B is a

[3] The original has no italics.

self-consistent observation-statement about a possible future event, and K any class of true observation-statement about past events, then B can *always* be conjoined with K without contradiction; or in other words, if we add a statement B about a possible future event to statements in K we can *never* arrive at a logical contradiction. Hume's finding can also be formulated as follows: *no logically possible future observation can ever contradict the class of past observations.*

Let us now add to Hume's simple finding a theorem of pure logic, namely: whenever a statement B can be conjoined without contradiction to a class of statements K, then it can also be conjoined without contradiction to any class of statement that consists of statements of K together with any statement that can be *derived from* K.

And so we have proved our point: if Newton's theory could be derived from a class K of true observation-statements, then no future observation B could possibly contradict Newton's theory and the observations K.

Yet it is known, on the other hand, that, from Newton's theory and past observations, we can logically derive a statement that tells us whether or not there will be an eclipse of the sun tomorrow. Now if this derived statement tells us that tomorrow there will be no eclipse of the sun, then our B is clearly *incompatible* with Newton's theory and the class K. From this and our previous results it follows logically that it is impossible to assume that Newton's theory can be derived from observations.

Thus we have proved our third point. And we can now see the whole riddle of experience — the paradox of the empirical sciences, as discovered by Kant:

Newton's dynamics goes essentially beyond all observations. It is universal, exact and abstract; it arose historically out of myths; and we can show by purely logical means that it is not derivable from observation-statements.

Kant also showed that what holds for Newtonian theory must hold for *everyday experience*, though not, perhaps, to quite the same extent: that everyday experience, too, goes far beyond all observation. Everyday experience too must *interpret* observation: for without theoretical interpretation, observation remains blind — uninformative. Everyday experience constantly operates with abstract ideas, such as that of cause and effect, and so it cannot be derived from observations.

In order to *solve* the riddle of experience, and to explain how natural science and experience are at all possible, Kant constructed his *theory of experience and of natural science*. I admire this theory as a truly heroic attempt to solve the paradox of experience, yet I believe

that it answers a false question, and hence that it is *in part* irrelevant. Kant, the great discoverer of the riddle of experience, was in error about one important point. But, his error, I hasten to add, was quite unavoidable, and it detracts in no way from his magnificent achievement.

What was this error? Kant, like almost all his contemporaries, and like almost all his successors right into the twentieth century, was convinced that Newton's theory was *true*. This conviction was inescapable. Newton's theory had made the most astonishing and exact predictions, all of which had proved strikingly correct. Only ignorant men could doubt its truth. How little we may reproach Kant for his belief is best shown by the fact that even Henri Poincaré, the greatest mathematician, physicist and philosopher of his generation, who died shortly before the First World War, believed like Kant that Newton's theory was true and irrefutable. Poincaré was one of the few scientists who felt about Kant's paradox almost as strongly as Kant himself; and though he proposed a solution which somewhat differed from Kant's, it was only a variant of it. The important point, however, is that he fully shared Kant's error, as I have called it. It was an unavoidable error — unavoidable, that is, before Einstein.

Even those who do not accept Einstein's theory of gravitation ought to admit that his was an achievement of truly epoch-making significance. For his theory established at least that Newton's theory, no matter whether true or false, was certainly *not the only possible* system of celestial mechanics that could explain the phenomena in a simple and convincing way. For the first time in more than 200 years, Newton's theory had become *problematical*. Newton's theory, that great masterpiece, had become, during these two centuries, a dangerous *dogma* — a dogma of almost stupefying power. I have no objection to those who oppose Einstein's theory on scientific grounds. But even Einstein's opponents, like his greatest admirers, ought to be grateful to him for having freed physics of that paralysing *idée fixe*, the belief in the incontestible truth of Newton's theory. We now look upon this theory, thanks to Einstein, as a hypothesis (or a system of hypotheses) — perhaps the most magnificent and the most important hypothesis in the history of science, and certainly an astonishing approximation to the truth.

Now if, unlike Kant, we consider Newton's theory as a hypothesis whose truth is problematic, then we must radically alter Kant's problem. No wonder, then, that his solution no longer suits the new post-Einsteinian formulation of the problem, and that it must be amended accordingly.

Kant's solution of the problem is well known. He assumed, correctly I think, that the world as we know it is our *interpretation* of observable facts: *that the world is our interpretation of these facts in the light of theories that we ourselves invent.* As Kant puts it: 'Our intellect does not draw its laws from nature . . . but imposes them upon nature.' While I consider this formulation of Kant's essentially correct, I feel that it is a little too radical, and I should therefore like to put it in the following modified form: 'Our intellect does not draw its laws from nature, but tries — with varying degrees of success — to impose upon nature laws which it freely invents.' The difference is this. Kant's formulation not only implies that our reason attempts to impose laws upon nature, but also that it is invariably successful in this. For Kant believed that Newton's laws were successfully imposed upon nature by us: that we were bound to interpret nature by means of these laws; from which he concluded that they must be true *a priori*. This is how Kant saw these matters; and Poincaré saw them in a similar way.

Yet we know since Einstein that very different theories and very different interpretations are also possible, and that they may even be superior to Newton's. Thus reason is capable of more than one interpretation. Nor can it impose its interpretation upon nature once and for all time. Reason works by trial and error. We invent our myths and our theories and we try them out: we try to see how far they take us. And we improve our theories if we can. The better theory is the one that has the greater explanatory power: that explains more; that explains with greater precision; and that allows us to make better predictions.

Since Kant believed that it was our task to explain the uniqueness and the truth of Newton's theory, he was led to the belief that this theory followed inescapably and with logical necessity from the laws of our understanding. The modification of Kant's solution which I propose, in accordance with the Einsteinian revolution, frees us from this compulsion. In this way, theories are seen to be the *free* creations of our own minds, the result of an almost poetic intuition, of an attempt to understand intuitively the laws of nature. But we no longer try to force our creations upon nature. On the contrary, we question nature, as Kant taught us to do; and we try to elicit from her *negative* answers concerning the truth of our theories: we do not try to prove or to *verify* them, but we test them by trying to *disprove* or to falsify them, to *refute* them.

In this way, the freedom and boldness of our theoretical creations can be controlled and tempered by self-criticism, and by the severest tests we can design. It is here, through our critical methods of

testing, that scientific rigour and logic enter into empirical science.

We have seen that theories cannot be logically derived from observations. They can, however, clash with observations: they can contradict observations. This fact makes it possible to infer from observations that a theory is *false*. The possibility of refuting theories by observations is the basis of all empirical tests. For the test of a theory is, like every rigorous examination, always an attempt to show that the candidate is mistaken — that is, that the theory entails a false assertion. From a logical point of view, all empirical tests are therefore *attempted refutations*.

In conclusion, I should like to say that, ever since Laplace, attempts have been made to attribute to our theories instead of *truth* at least *a high degree of probability*. I regard these attempts as misconceived. All we can ever hope to say of a theory is that it explains this or that; that it has been tested severely, and that it has stood up to all our tests. We may also compare, say, two theories in order to see which of them has stood up better to our severest tests — or in other words, which of them is *better corroborated* by the results of our tests. But it can be shown by purely mathematical means that *degree of corroboration can never be equated with mathematical probability*. It can even be shown that all theories, including the best, have the same probability, namely zero. But the degree to which they are corroborated (which in theory at least, can be found out with the help of the calculus of probability) may approach very closely to unity, i.e. its maximum, though the probability of the theory is zero. That an appeal to probability is incapable of solving the riddle of experience is a conclusion first reached long ago by David Hume.

Thus logical analysis shows that experience does not consist in the mechanical accumulation of observations. Experience is creative. It is the result of free, bold and creative interpretations, controlled by severe criticism, and by severe tests.

II. THE PROBLEM OF THE IRREFUTABILITY OF PHILOSOPHICAL THEORIES

In order to avoid right from the start the danger of getting lost in generalities, it might be best to explain at once, with the help of five examples, what I mean by a *philosophical theory*.

A typical example of a philosophical theory is Kant's doctrine of *determinism*, with respect to the world of experience. Though Kant was an indeterminist, he asserted in the *Critique of Practical Reason*[4] that adequate knowledge of our psychological and physiological conditions and of our environment would make it possible to predict

[4] *Kritik der praktischen Vernunft*, 4th to 6th ed., p. 172; *Works*, ed. Cassierer, vol. 5, p. 108.

our future behaviour with the same certainty with which we can predict an eclipse of the sun or of the moon.

In more general terms, one could formulate the determinist doctrine as follows.

The future of the empirical world (or of the phenomenal world) is completely predetermined by its present state, down to its smallest detail.

Another philosophic theory is *idealism*, in Berkeley's or Schopenhauer's formulation; we may express it by the following thesis: 'The empirical world is my idea', or '*The world is my dream*'.

A third philosophic theory — and one that is very important today — is epistemological *irrationalism*, which might perhaps be explained as follows.

Since we know from Kant that human reason is incapable of grasping, or knowing, the world of things in themselves, we must either give up the hope of ever knowing it, or else try to know it otherwise than by means of our reason; and since we cannot and will not give up this hope, we can only use irrational or supra-rational means, such as instinct, poetic inspiration, moods, or emotions.

This, irrationalists claim, is possible because in the last analysis, we are ourselves such things-in-themselves; thus if we can manage somehow to obtain an intimate and immediate knowledge of ourselves, we can thereby find out what things-in-themselves are like.

This simple argument of irrationalism is highly characteristic of most nineteenth-century post-Kantian philosophers; for example, of the ingenious Schopenhauer who in this way discovered that since we, as things-in-themselves, are *will*, will must be the thing-in-itself. The world, as a thing-in-itself, is *will*, while the world as phenomenon is an *idea*. Strangely enough, this obsolescent philosophy, dressed up in new clothes, has once again become the latest fashion, although, or perhaps just because, its striking similarity to old post-Kantian ideas has remained hidden (so far as anything may remain hidden under the emperor's new clothes). Schopenhauer's philosophy is nowadays propounded in obscure and impressive language, and his self-revealing intuition that man, as a thing-in-itself, is ultimately *will*, has now given place to the self-revealing intuition that man may so utterly bore himself that his very boredom proves that the thing-in-itself is Nothing — that it is Nothingness, Emptiness-in-itself. I do not wish to deny a certain measure of originality to this existentialist variant of Schopenhauer's philosophy: its originality is proved by the fact that Schopenhauer could never have thought so poorly of his powers of self-entertainment. What he discovered in himself was will, activity, tension, excitement — roughly the opposite of what some existentialists discovered: the utter boredom of the

bore-in-himself bored by himself. Yet Schopenhauer is no longer the fashion: the great fashion of our post-Kantian and post-rationalist era is what Nietzsche ('haunted by premonitions, and suspicious of his own progeny') rightly called 'European nihilism'.[5]

Yet all this is only by the way. We now have before us a list of five philosophical theories.

First, determinism: the future is contained in the present, inasmuch as it is fully determined by the present.

Second, idealism: the world is my dream.

Third, irrationalism: we have irrational or suprarational experiences in which we experience ourselves as things-in-themselves; and so we have some kind of knowledge of things-in-themselves.

Fourth, voluntarism: in our own volition we know ourselves as will. The thing-in-itself is the will.

Fifth, nihilism: in our boredom we know ourselves as nothings. The thing in itself is Nothingness.

So much for our list. I have chosen my examples in such a way that I can say of each one of these five theories, after careful consideration, that I am convinced that it is false. To put it more precisely; I am *first* of all an indeterminist, *secondly* a realist, *thirdly* a rationalist. As regards my fourth and fifth examples, I gladly admit — with Kant and other critical rationalists — that we cannot possess anything like full knowledge of the real world with its infinite richness and beauty. Neither physics nor any other science can help us to this end. Yet I am sure the voluntarist formula, 'The world is will', cannot help us either. And as to our nihilists and existentialists that bore themselves (and perhaps others), I can only pity them. They must be blind and deaf, poor things, for they speak of the world like a blind man of Perugino's colours or a deaf man of Mozart's music.

Why, then, have I made a point of selecting for my examples a number of philosophical theories that I believe to be false? Because I hope in this way to put more clearly the problem contained in the following important statement.

Although I consider each one of these five theories to be *false*, I am nevertheless convinced that each of them is *irrefutable*.

With this statement I have at last arrived at our problem. The statement is so important that I wish to repeat it:

'Although I consider each one of these five theories to be *false*, I am nevertheless convinced that each of them is *irrefutable*.'

[5] Cf. Julius Kraft, *Von Husserl zu Heidegger*, 2nd ed. 1957, e.g. pp. 103f, 136f and particularly p. 130, where Kraft writes: 'Thus it is hard to understand how existentialism could ever have been considered to be something new in philosophy, from an epistemological point of view.' Cf. also H. Tint, *Proc. Aristot. Society*, 1956-7, pp. 253ff.

Listening to this statement, you may well wonder how I can possibly hold a theory to be *false and irrefutable* at one and the same time – I who claim to be a rationalist. For how can a rationalist say of a theory that it is false and irrefutable? Is he not bound, as a rationalist, to refute a theory before he asserts that it is false? And conversely, is he not bound to admit that if a theory is irrefutable, it is true?

The last question can be answered fairly simply. There have been thinkers who believed that the truth of a theory may be inferred from its irrefutability. Yet this is an obvious mistake, considering that there may be two incompatible theories which are equally irrefutable – for example, determinism and its opposite, indeterminism. Now since two incompatible theories cannot both be true, we see from the fact that both theories are irrefutable that irrefutability cannot entail truth.

To infer the truth of a theory from its irrefutability is therefore inadmissible, no matter how we interpret irrefutability. For normally 'irrefutability' would be used in the following two senses:

The first is a purely logical sense: we may use 'irrefutable' to mean the same as 'irrefutable by purely logical means'. But this would mean the same as 'consistent'. Now it is quite obvious that the truth of a theory cannot possibly be inferred from its consistency.

The second sense of 'irrefutable' refers to refutations that make use not only of logical (or analytic) but also of empirical (or synthetic) assumptions; in other words, it admits empirical refutations. In this second sense, 'irrefutable' means the same as 'not empirically refutable', or more precisely 'compatible with any possible empirical statement' or 'compatible with every possible experience'.

Now both the logical and the empirical irrefutability of a statement or a theory can easily be reconciled with its falsehood. In the case of logical irrefutability this is clear from the fact that every empirical statement and its negation must both be *logically* irrefutable. For example, the two statements 'Today is Monday' and 'Today is not Monday' are both logically irrefutable; but from this it follows immediately that there exist false statements which are logically irrefutable.

Regarding empirical irrefutability, the situation is a little different. The simplest examples of empirically irrefutable statements are so-called strict or pure existential statements. An example of a strict or pure existential statement would be the following. 'There exists a pearl which is ten times larger than the next largest pearl.' If in this statement we restrict the words 'There exists' to some finite region in space and time, then it may of course become a refutable

statement. For example, the following statement is obviously empirically refutable: 'At this moment and in this box here there exist at least two pearls one of which is ten times larger than the next largest pearl in this box.' But then this statement is no longer a strict or pure existential statement; rather it is a *restricted* existential statement. A strict or pure existential statement applies to the whole universe, and it is irrefutable simply because there can be no method by which it could be refuted. For even if we were able to search our entire universe, the strict or pure existential statement would not be refuted by our failure to discover the required pearl, seeing that it might always be hiding in a place where we are not looking.

Examples of empirically irrefutable existential statements which are of greater interest are the following.

'There exists a completely effective cure for cancer, or more precisely, there is a chemical compound which can be taken without ill effect, and which cures cancer.' Needless to say, this statement must not be interpreted as meaning that such a chemical compound is actually *known* or that it will be discovered within a given time.

Similar examples are: 'There exists a cure for any infectious disease' and 'There exists a Latin formula which, if pronounced in proper ritual manner, cures all diseases.'

Here we have an empirically irrefutable statement that few of us would hold to be true. The statement is irrefutable because it is obviously impossible to try out every conceivable Latin formula in combination with every conceivable manner of pronouncing it. Thus there always remains the logical possibility that there might be, after all, a magical Latin formula with the power of curing all diseases.

Even so, we are justified in believing that this irrefutable existential statement is false. We certainly cannot *prove* its falsehood; but everything we know about diseases tells against it being true. In other words, though we cannot establish its falsity, the conjecture that there is no such magical Latin formula is much more reasonable than the irrefutable conjecture that such a formula does exist.

I need hardly add that through almost 2000 years learned men have believed in the truth of an existential statement very much like this one; this is why they persisted in their search for the philosopher's stone. Their failure to find it does not prove anything — precisely because existential propositions are irrefutable.

Thus the logical or empirical irrefutability of a theory is certainly not a sufficient reason for holding the theory to be true, and hence I have vindicated my right to believe at the same time that these five philosophical theories are irrefutable, and that they are false.

Some twenty-five years ago, I proposed to distinguish empirical or scientific theories from non-empirical or non-scientific ones precisely by defining the empirical theories as the refutable ones and the non-empirical theories as the irrefutable ones. My reasons for this proposal were as follows. Every serious test of a theory is an attempt to refute it. Testability is therefore the same as refutability, or falsifiability. And since we should call 'empirical' or 'scientific' only such theories as can be empirically tested, we may conclude that it is the possibility of an empirical refutation which distinguishes empirical or scientific theories.

If this 'criterion of refutability' is accepted, then we see at once that *philosophical* theories will be *irrefutable by definition.*

My assertion that our five philosophical theories are irrefutable may now sound almost trivial. At the same time it will have become obvious that, though I am a rationalist, I am in no way obliged to refute these theories before being entitled to call them 'false'. And this brings us to the crux of our problem:

If philosophical theories are all irrefutable, how can we ever distinguish between true and false philosophical theories?

This is the serious problem which arises from the *irrefutability of philosophical theories.*

In order to state the problem more clearly, I should like to reformulate it as follows.

We may distinguish here between three types of theory.

First, logical and mathematical theories.

Secondly, empirical and scientific theories.

Thirdly, philosophical theories.

How, in each of these groups, can we distinguish between true and false theories?

Regarding the first group the situation is clear. Whenever we find a mathematical theory of which we do not know whether it is true or false, we test it, first superficially and then more severely, by trying to refute it. If we are unsuccessful, we then try to prove it or to refute its negation. If we fail again, doubts as to the truth of the theory may have cropped up again, and we shall again try to refute it, and so on, until we either reach a decision or else shelve the problem as too difficult for us.

The situation could also be described as follows. Our task is the testing, the critical examination, of two (or more) rival theories. We solve it by trying to refute them — either the one or the other — until we come to a decision. In mathematics (but only in mathematics) such decisions are generally *final*: invalid proofs that escape detection are rare.

If we now look at the empirical sciences, then we find that we follow, as a rule, fundamentally the same procedure. Once again we test our theories: we examine them critically, we try to refute them. The only important difference is that now we can also make use of empirical arguments in our critical examinations. But these empirical arguments occur only together with other critical considerations. Critical thought as such remains our main instrument. Observations are used only if they fit into our critical discussion.

Now if we apply these considerations to philosophical theories, our problem can be reformulated as follows:

Is it possible to examine irrefutable philosophical theories *critically*? If so, what can a critical discussion of a theory consist of, if not of *attempts to refute the theory*?

In other words, is it possible to assess an irrefutable theory rationally — which is to say, critically? And what reasonable argument can we adduce for and against a theory of which we know that it is neither demonstrable nor refutable?

In order to illustrate these various formulations of our problem by examples, we may first refer again to the problem of determinism. Kant knew perfectly well that we are unable to predict the future actions of a human being as accurately as we can predict an eclipse. But he explained the difference by assuming that we know far less about the present conditions of a man — about his wishes and fears, his feelings and his motives — than about the present state of the solar system. Now this assumption contains, implicitly, the following hypothesis:

'There exists a true description of the present state of this man which would suffice (in conjunction with true natural laws) for the prediction of his future actions.' This of course is again a purely existential statement and thus irrefutable. Can we, in spite of this fact, discuss Kant's argument rationally and critically?

As a second example we may consider the thesis: 'The world is my dream.' Though this thesis is clearly irrefutable, few will believe in its truth. But can we discuss it rationally and critically? Is not its irrefutability an insurmountable obstacle to any critical discussion?

As to Kant's doctrine of determinism, we might perhaps think that the critical discussion of it might begin by saying to him: 'My dear Kant, it simply is not enough to assert that there exists a true description that is sufficiently detailed to enable us to predict the future. What you must do is tell us exactly what this description would consist of, so that we may test your theory empirically.' This speech, however, would be tantamount to the assumption that philosophical — that is, irrefutable — theories can never be discussed

and that a responsible thinker *is bound* to replace them by empirically testable theories, in order to make a rational discussion possible.

I hope that our *problem* has by now become sufficiently clear; so I will now proceed to *propose a solution* of it.

My solution is this: if a philosophical theory were no more than an isolated assertion about the world, flung at us with an implied 'take it or leave it' and without a hint of any connection with anything else, then it would indeed be beyond discussion. But the same might be said of an empirical theory also. Should anybody present us with Newton's equations, or even with his arguments, without explaining to us first what the problems were which his theory was meant to solve, then we should not be able to discuss its truth rationally — no more than the truth of the *Book of Revelation*. Without any knowledge of the results of Galileo and Kepler, of the problems that were resolved by these results, and of Newton's problem of explaining Galileo's and Kepler's solutions by a unified theory, we should find Newton's theory just as much beyond discussion as any metaphysical theory. In other words every *rational* theory, no matter whether scientific or philosophical, is rational in so far as it tries to *solve certain problems*. A theory is comprehensible and reasonable only in its relation to a given *problem-situation*, and it can be rationally discussed only by discussing this relation.

Now if we look upon a theory as a proposed solution to a set of problems, then the theory immediately lends itself to critical discussion — even if it is non-empirical and irrefutable. For we can now ask questions such as, Does it solve the problem? Does it solve it better than other theories? Has it perhaps merely shifted the problem? Is the solution simple? Is it fruitful? Does it perhaps contradict other philosophical theories needed for solving other problems?

Questions of this kind show that a critical discussion even of irrefutable theories may well be possible.

Once again, let me refer to a specific example: the idealism of Berkeley or Hume (which I have replaced by the simplified formula 'The world is my dream'). It is notable that these authors were far from wishing to offer us so extravagant a theory. This may be seen from their repeated insistence that their theories were really in agreement with sound common sense. And if we try to understand the *problem situation* which induced them to propound this theory, then we find that Berkeley and Hume believed that all our knowledge was reducible to *sense-impressions* and to associations between *memory-images*. This assumption led these two philosophers to adopt idealism; and in the case of Hume, in particular, very unwillingly. Hume was

an idealist only so far as he failed in his attempt to reduce realism to *sense-impressions*.

It is therefore perfectly *reasonable* to criticize Hume's idealism by pointing out that his sensualistic theory of knowledge and of learning was in any case inadequate, and that there are less inadequate theories of learning which have no unwanted idealistic consequences.

In a similar way, we could now proceed to discuss Kant's determinism rationally and critically. Kant was, in his fundamental intention, an indeterminist: even though he believed in determinism with respect to the phenomenal world as an unavoidable consequence of Newton's theory, he never doubted that man, as a moral being, was not determined. Kant never succeeded in solving the resulting conflict between his theoretical and practical philosophy in a way that satisfied himself completely, and he despaired of ever finding a real solution.

Now in the setting of this *problem-situation*, it becomes possible to criticize Kant's determinism. We may ask, for example, whether it really follows from Newton's theory. Let us conjecture for a moment that it does not. I do not doubt that a clear proof of the truth of this conjecture would have persuaded Kant to renounce his doctrine of determinism – even though this doctrine happens to be irrefutable and even though he would not, for this very reason, have been logically compelled to renounce it.

Similarly with irrationalism. It first entered rational philosophy with Hume – and those who have read Hume, that calm analyst, cannot doubt that this was not what he intended. Irrationalism was the unintended consequence of Hume's conviction that *we do in fact learn* by Baconian induction, coupled with Hume's logical proof that *it is impossible rationally to justify induction*. 'So much the worse for rational justification' was a conclusion which Hume, of necessity, was compelled to draw from this situation. He accepted this irrational conclusion with the integrity characteristic of the real rationalist who does not shrink from an unpleasant conclusion if it seems to him unavoidable.

Yet in this case, it was not unavoidable, though it seemed to be so to Hume. We are not in fact the Baconian induction machines that Hume believed us to be. Habit or custom does not play the role in the process of learning which Hume assigned to it. And so Hume's problem dissolves and with it his irrationalist conclusions.

The situation of post-Kantian irrationalism is somewhat similar. Schopenhauer, in particular, was genuinely opposed to irrationalism. He wrote with only *one* desire: to be understood; and he wrote more lucidly than any other German philosopher. His striving to be

understood made him one of the few great masters of the German language.

Yet Schopenhauer's problems were those of Kant's metaphysics — the problem of determinism in the phenomenal world, the problem of the thing in itself, and the problem of our own membership of a world of things in themselves. He solved these problems — *problems transcending all possible experience* — in his typically rational manner. But the solution was *bound* to be irrational. For Schopenhauer was a Kantian and as such he believed in the Kantian limits of reason: he believed that the limits of human reason coincided with *the limits of possible experience*.

But here again, there are other possible solutions. Kant's problems can and must be revised; and the direction that this revision should take is indicated by his fundamental idea of critical, or self-critical, rationalism. The discovery of a philosophical problem can be something final; it is made once, and for all time. But the solution of a philosophical problem is never final. It cannot be based upon a final proof or upon a final refutation: this is a consequence of the irrefutability of philosophical theories. Nor can the solution be based upon the magical formulas of inspired (or bored) philosophical prophets. Yet it may be based upon the conscientious and critical examination of a problem-situation and its underlying assumptions, and of the various possible ways of resolving it.

THE PHILOSOPHY OF R. G. COLLINGWOOD

By E. W. F. Tomlin

I

In the concluding chapter of his *Essay on Philosophical Method*, Collingwood discusses the subject of 'Philosophy as a branch of literature'. He begins by examining the relations between philosophy and poetry. 'A philosophical work', he writes, 'if it must be called a poem, is not a mere poem, but a poem of the intellect. What is expressed in it is not emotions, desires, feelings as such, but those which a thinking mind experiences in its search for knowledge; and it expresses these only because the experience of them is an integral part of the search, and that search is thought itself. When this qualification is added, it becomes plain that philosophical literature is in fact prose; it is poetry only in the sense in which all prose is poetry — poetry modified by the expression of a content, something which the writer is trying to say.'

This passage, so characteristic of Collingwood's middle period, reveals the pervasive influence of Croce. In the final sentence, pure poetry would seem to be defined as expression without content. This is a view which Croce in his *Essence of Aesthetic* (1912) modified, and to which Collingwood in *The Principles of Art* (1937) no longer subscribed. But the passage contains a truth which Collingwood consistently practised. Despite the difficult conditions in which most of his books were written (he was dogged by ill-health), they form literary masterpieces. His prose is firm, supple, often glittering, always clear. In the same chapter, he points out how 'an imaginative and somewhat poetic style', though out of place in science and history, is in philosophy often highly successful. Finally, he clinches his argument by affirming that 'the philosopher must go to school with the poets in order to learn the use of language'.

The Essay on Philosophical Method — and it is interesting that a treatise on *method* should contain a chapter on this subject — was published in 1932. In his later *Autobiography* (1939), Collingwood described it as 'my best book in matter: in style I may call it my only book, for it is the only one I ever had time to finish as well as I knew how, instead of leaving it in a more or less rough state'. To many students at Oxford, the *Essay* proved a landmark. It appeared at a crucial moment. The Grand Old Men of Oxford philosophy, Bradley and Bosanquet, no longer satisfied. The succeeding genera-

tion of Idealists, led by J. A. Smith and Joachim (the latter a close friend of Bradley), had committed little to paper. The present writer, who recalls vividly the farewell lecture of Joachim, felt that with the passing of these men an era had ended. It was an era characterized by a certain manner as well as by a prescribed matter: Joachim read fastidiously and with exquisite diction from a hand-written document held close to the eyes. Meanwhile, a group hardly out of its teens had invaded the lecture-rooms. Of these the most vocal and brilliant was A. J. Ayer, whose *Language, Truth and Logic* appeared in 1933. In the estimation of some, including its author, this book swept away not merely Idealism but philosophy itself.

The clamour raised by the Logical Positivists has never really subsided. The title came to stand for a point of view, almost a way of life. As such, it generated and indeed implied a metaphysics of its own. This has been recognized even by those in close sympathy with its anti-metaphysical tendencies.[1] Today, not because it is outmoded but because its metaphysics has permeated the atmosphere of thought, Logical Positivism is a label which British thinkers are inclined to avoid. It has suffered the fate of that other label, Existentialism, which Continental philosophers, from Marcel to Camus, are now quick to disown. Nevertheless, while repudiating orthodox Logicial Positivism, thinkers such as Russell and Ayer still acknowledge that it comes nearest to representing their general point of view. Those who prefer to describe themselves as Analytical Philosophers claim to have originated a form of philosophy different in *kind* from that of the past. They even claim to have discovered for the first time what philosophy is.[2]

Collingwood is a figure of interest for three main reasons. In the first place, he remained even in his final phase a traditional philosopher. To him philosophy was an adventure in thought. The philosopher engaged in his task in order to find out something and, in so doing, to become something. In short, philosophy was committed to the unity of theory and practice. Secondly, and because of this, Collingwood formed a link with the generation of Bradley and Bosanquet. His succession to J. A. Smith as Waynflete Professor of Metaphysics at Oxford — at the latter's wish, it is reported — continued a form of Apostolic succession broken only by his resignation in 1942. Consequently, he is still classified as an Idealist, by those who find meaning and utility in such labels. Thirdly, by

[1] Vide *Philosophy and Analysis* (1954), edited with an introduction by Margaret Macdonald, p. 9, and 'Logical Translation' by Isaiah Berlin (*Proceedings of the Aristotelian Society*, June 5th, 1950, p. 188).

[2] Vide *The Nature of Metaphysics* (1957), edited by D. F. Pears, chapter on 'Metaphysics and Ethics' by Iris Murdoch, p. 123.

reason of his keen intellect, his linguistic ability, and no doubt some inherited predisposition, he was in closer touch with philosophical movements abroad than most of his contemporaries. And he had wider and more practical interests.

These reasons combined to make him not merely a significant but an isolated figure. He was aloof, because preoccupied. At his untimely death in 1943 his name received less than its due regard. His last books, though among his most challenging, failed to excite critical attention: *The Principles of Art*, the best-written work of his later period, was ignored by *Mind*. Taken as a whole, his achievement has never received the attention bestowed upon less systematic, if superficially more brilliant, work. No mention is made of him in John Laird's *Recent Philosophy* (1936), though Laird later made reparation in a review in *Philosophy* of *The New Leviathan* (1942). Only Guido de Ruggiero's *Filosofi del Novecento* (1931) did him justice, but at that time Collingwood's philosophical works were limited to two volumes of philosophy and a brief essay on aesthetic theory. A little monograph by the present writer met with success, not so much because of its intrinsic merit, as on account of its being the only general study of Collingwood outside the specialist periodicals.[3]

After some years of comparative neglect, at least by British philosophers, there are signs of revived interest in Collingwood's work. This is a welcome development. Recognition, however critical or uncritical it may be, has come from unexpected quarters. The chapter on Causation in the *Essay on Metaphysics* has exerted influence upon legal opinion, especially in the United States. Volume IX of Arnold Toynbee's *Study of History* contains an interesting, if highly critical, analysis of his ideas on historiography. Ex-pupils have spoken of him on the Third Programme of the B.B.C. The *Autobiography* appeared in the famous Penguin series. He has been attacked — and defended — in the pages of *The New Statesman*.

Nevertheless, the Collingwood revival, if such it can be called, is not so wholehearted as one might have expected or hoped. With certain exceptions, his colleagues at Oxford have been grudging in their praise, unsparing in their blame: witness the respectful but inadequate account of his work by H. J. Paton.[4] For this there are perhaps two reasons. In the first place, Collingwood invited neither praise nor blame for his work: he invited disinterested attention to the problems he was studying. He rarely engaged in public controversy. There are few references in his books or essays to

[3] *R. G. Collingwood* (Longman's, 1953).
[4] 'Fifty years of Philosophy' (*Contemporary British Philosophy 1956*, p. 345).

immediate or personal issues. A certain disdain, and in his later books something like truculence, may have caused offence. 'I take this opportunity', he wrote in his *Autobiography*, 'of saying that I will not be drawn into discussion of what I write. Some readers may wish to convince me that it is all nonsense. I know how they would do it; I could invent their criticisms for myself. Some may wish to show me that on this or that detail I am wrong. Perhaps I am: if they are in a position to prove it, let them write not about me but about the subject, showing that they can write about it better than I can; and I will read them gladly. And if there are any who think my work good, let them show their approval of it by attention to their own.' In the second place, the problems with which Collingwood was preoccupied were peculiar to himself. The movement in British thought initiated by G. E. Moore's famous essay *The Refutation of Idealism* (1905) was largely a movement of liquidation. To the next generations, the task of philosophy was almost exclusively critical. To what lengths the critical process could be pushed may be judged by an interesting fact: the two philosophical 'best-sellers' during the last twenty years have been Ayer's first book and Ryle's *Concept of Mind*, both iconoclastic documents. An orthodoxy of scepticism has been installed. There is distrust of positive affirmations. In a recent essay, Ryle has publicly affirmed that the time is not ripe for syntheses or systems.[5] An interest in problems of historiography, cosmology, ethics and political theory has come to be regarded as an extra-philosophical luxury. They were Collingwood's interests.

To lay the blame for the vagaries of Collingwood's reputation solely upon his critics would be unfair. There are certain intractable elements in his thought which, present from the beginning, worked themselves gradually to the surface. At some point he allowed his intense interest in history and historical investigation so to dominate his mind as to warp his judgment. Take his posthumous volume *The Idea of History* (1945). The book falls into two parts (indeed he intended to work them up into separate volumes): first there is a brilliant account of past historiography, and then there follows a series of essays of varying length and quality. Although his private papers testify to the high hopes he placed upon the book, the impression given is that of an immense prolegomenon. It was a work for the completion of which he lacked a guiding thread. The greater part of *The Idea of History* consists of lectures delivered at Oxford in 1936. As the present writer can confirm from his personal notes, the lectures underwent little fundamental revision. Yet Collingwood found the time and energy between 1936 and the year of his

[5] *The Nature of Metaphysics* ('Final Discussion'), p. 156.

death, 1943, to complete four other books: *The Principles of Art* (1937), the *Autobiography* (1939) and *Essay on Metaphysics* (1940) and *The New Leviathan* (1942). Moreover, he left in manuscript a carefully written if incomplete work, *The Idea of Nature* (1946). Of these books *The New Leviathan* is perhaps the most systematic; *The Principles of Art*, forming the most sustained piece of prose after the *Essay on Philosophical Method*, runs it close. The *Autobiography* and the *Essay on Metaphysics*, fine though they are in places, seem to have been written by a man aware of the approach of the 'night when no man can work'. While racing to complete them, Collingwood must have dwelt upon the work by which he hoped to establish his reputation. Why did he not concentrate upon it? We may hazard an answer. To write such a book he needed a 'metaphysics': and in the *Essay on Metaphysics* he had destroyed — it must be confessed with some plausibility — the constructive system on which alone such a masterpiece could have been based.

I therefore raise early in this essay and without apology the vexed question of Collingwood's 'historicism'. The idea that metaphysics forms a branch of history, that is to say the 'history of history', was outlined first in his inaugural lecture *The Historical Imagination* (1935). In this paper Collingwood first moved away from the idea that metaphysics is the Science of the Absolute, as the *Essay on Philosophical Method* had affirmed in an interesting passage (Chapter VI, § 2) and as Collingwood shows evidence of having believed as late as 1936. Yet he had every reason not to abandon the traditional view. Of his successive analyses of the attitude of past philosophers to metaphysics, the passage in the *Essay on Philosophical Method* on Kant (Introduction, x) is perhaps the most compelling. Given the fact that most books designed to throw light on *The Critique of Pure Reason* greatly exceed that book in obscurity (as the student realizes when, having been fed upon handbooks, he nervously approaches the original text), Collingwood's brief elucidation is masterly. His case against Kant is as follows. Far from seeking to discredit metaphysics, Kant wished to clear away obstacles to its right pursuit. The *Critique* formed 'a propaedeutic or introduction to metaphysics'. After mastering this propaedeutic, Kant hoped that the philosopher '*would go back to his proper work, that of metaphysical speculation*' (my italics). In fact the propaedeutic raised so many problems and absorbed so much of Kant's own attention, that it became an end instead of a means. To that extent, the shadow of Kant is still upon us. Yet problems of methodology, as Collingwood points out, cannot be divorced from those of substantive philosophy. The study of philosophy, if consistently pursued, raises at every stage the problems

of methodology. One activity cannot assume precedence over the other. (The *Essay*, for instance, is interspersed with philosophical reflections.) In short, philosophy has no propaedeutic. There is no preparation for thought. Having elaborated a method, the philosopher cannot go back to his 'proper work': he can only go on with it. But this is impossible if his so-called propaedeutic, as in the case of that which Collingwood elaborated in his *Essay on Metaphysics*, is not an introduction at all but a demonstration that metaphysics is not metaphysics but something else, namely history.

Secondly, Collingwood's sympathy and friendship with Croce, particularly the Croce of the *Theory of History and Historiography*, might have warned him of the obvious dangers of historicism. Croce is a thinker for whom British philosophers, save a few such as Bosanquet and J. A. Smith, have shown little affection. This may be due to a style difficult to render in English otherwise than as a kind of impassioned prose-poetry. (The volume *My Philosophy*, edited and translated by E. F. Carritt in 1945, presented Croce in much better light.) Nevertheless, Croce, a formidable analyst and himself a skilled historian, is at his most masterly in his attacks upon historicism. That a work of such powerful influence upon Collingwood as the *Theory of History* should have failed to rescue him from this pitfall is all the more strange. The reduction of metaphysical speculation to an historical inquiry into what presuppositions 'happen to have been held' at a particular time is a lame conclusion to Collingwood's search for order and system. Moreover, there was the eloquent peroration of the *Essay on Philosophical Method* to furnish another warning. Here Collingwood observed that 'to a person who does not understand what philosophy is, or by what processes it moves, the history of these sixty generations appears as a chaos, the record of random movements hither and thither by wandering planets, which no theory of epicycles can reduce to reason. But this appearance of irrationality, I make bold to say, cannot survive the discovery that thought has a structure of its own, and the hypothesis that in its changes it is obeying the laws of that structure. Thus from the point of view of a rational system of philosophy, the past history of philosophical thought no longer appears as irrational: it is a body of experience to which we can appeal with confidence, because we understand the principles at work in it, and in the light of these principles find it intelligible'. In the *Essay on Metaphysics*, all that remains of the 'laws of structure' and the 'principles at work' are certain 'strains' or 'pressures' in society. These strains are the means whereby one constellation of presuppositions gives way to another. Of the nature and operation of such 'strains' we know little.

Consequently, we do not know whether it is desirable that, having made their appearance, they should be the means of transforming the metaphysical climate. Nor, in view of Collingwood's references to 'wandering planets', can it be said that the 'constellations' are any better grouped or organized; they are merely provisional equilibria. How can we 'appeal with confidence' to that which, on Collingwood's own admission, is merely a group of assumptions characteristic of a certain epoch — the 'thoughts that philosophers have had about it'? Confidence, like faith of any sort, is in that which is presumed true; how else would it lay claim to our fidelity? Yet to Collingwood the presuppositions of metaphysics, being the basis of all suppositions, are neither true nor false.

II

Unlike the practitioner of the arts, the philosopher is often judged upon the sum-total of his work. It is assumed that what he writes at the end of his career must represent a logical development, a necessary continuation, of that written earlier. If the final phase is seen to be fallacious, the critic is inclined to dismiss the whole as suspect. Such an assumption, which seems to be of comparatively recent growth, is to be deplored. Just as it would be foolish to judge Wordsworth on the basis of the inferiority of the *Ecclesiastical Sonnets* and parts of *The Excursion*, so it would be foolish — and in fact it has never been done — to judge Berkeley on *Siris* and *The Querist*, which, though not inferior to the early works, are different in both quality and outlook. Admittedly, Collingwood's thought ran off the rails in certain, though not all, of his last books: but it is surely mistaken to assume that because of a few such aberrations, his total achievement may be impugned. Bradley has not been expelled from the philosophical pantheon because of the fundamental inconsistency in *Appearance and Reality* between his denial of the category of relation and his assertion that all appearance is *of* reality. Despite their blemishes, Collingwood's last works are of great interest and importance. As a devoted friend, Professor T. A. Knox may have tended out of scrupulous regard for fairness to play down the merits of some of them.[6] Yet *The New Leviathan* alone would have assured a man's reputation; and to what recent works on natural science and aesthetics respectively can *The Idea of Nature* and *The Principles of Art* be considered inferior? The high standard Collingwood set himself has been responsible, as much as anything, for his immediate posthumous denigration. If a man fails to live up to our own standards, we

[6] *The Idea of History*, Editor's Preface, p. xx.

extend him sympathy. If he fails to live up to his own standards, we are reluctant to forgive him.

The emphasis upon Collingwood's final illness — a form of brain disease which led to total paralysis — can mislead. He confessed in the *Autobiography* never to have 'known any illness interfere with my power of thinking and writing, or with the quality of what I think and write'. Indeed, without his illnesses, Collingwood might never have written the great bulk of his work. At both ends of his career, the sense of urgency and anxiety possessed him. Early in the First World War, he published his first book, *Religion and Philosophy*. The book is now forgotten, superseded; but to Collingwood, writing at 'a time when a young man's expectation of life was a rapidly dwindling asset', it was the means of expounding his whole thought. The first book of a young thinker is also somehow his last. Into it he puts everything. With maturity comes the awareness that books form interim *communiqués*, and that the life of thought is never finished, only interrupted. Despite terminology evocative of Bradley and Bosanquet, *Religion and Philosophy* is worth reading, if the student can lay hands on a copy. Collingwood's object is to demonstrate that religion is not emotion but knowledge. This is a point to which he adhered; it is revived in *The Idea of History*.[7] Rejecting the psychological view of religion of William James's *Varieties of Religious Experience*, he went too far in a contrary direction. He proved too much. Religion, as he later recognized, had become too intellectual. In becoming intellectualized, it merged with theology and philosophy. *Speculum Mentis*, published in 1924, restored to religion its unique character. Religion is not knowledge as a whole; it is knowledge at one of its phases or rather crises. Whereas art is knowledge under the guise of imagination — the 'sleep of the soul', as he describes it — religion is knowledge at its first awakening. It is also feeling; but feeling may be knowledge in so far as it assumes direction. The religious phase of knowledge is the phase at which implicit knowledge first becomes explicit. Consequently, it is with the assertion of the religious impulse that man's personality first blossoms forth in creativity. Art may be logically anterior to religion, but it is at the religious stage that art receives its first real impetus.

To the reader coming fresh to Collingwood's work, the perfection of *Speculum Mentis* startles and exhilarates. In both form and content it is a remarkable achievement. Looking through some of the reviews it elicited, including those in *Philosophy* and *Mind*, one is surprised at their comparative coldness. Respect is grudging. One reviewer dismissed the book in a few lines as 'the usual idealistic nonsense'.

[7] 'Epilegomena', Ch. v, p. 314.

It is difficult to know precisely how much during thirty years it has been read. Few works appear to be in its debt: Michael Oakeshott's *Experience and its Modes*, another neglected work, is an exception. In professional circles one suspects that it fell upon deaf ears. Having written it, Collingwood seems to have turned his back upon synthesis, though a chapter in the *Essay on Philosophical Method* defends the notion that all philosophies implicitly recognize a system. In his *Autobiography* (Chapter VII), he referred to *Speculum Mentis* somewhat contemptuously. Then, having re-read it, he withdrew most of his strictures in a footnote.

Speculum Mentis did not spring ready-fashioned from its author's brain. It was a logical development of *Religion and Philosophy*. The later work grew out of a recognition of the shortcomings of the earlier.[8] In tone it is a post-war book, a call to order. For our generation its equivalent are the works of Sartre and Merleau-Ponty on the Continent and those of Ryle and Wisdom in Britain. To say as much is to show to what extent the second catastrophe has borne out some of Collingwood's own forebodings after the first; for with all their disciplined competence, the works of these writers are chiefly essays in demolition and mental purgation. Apart from an athletic style and a learning lightly carried, *Speculum Mentis* is distinguished by a sense of earnestness and responsibility. Son of an artist, Collingwood remained throughout life a practitioner of the arts; but he was concerned not merely with the 'fate of the artist in the modern world', to employ a familiar trope, but with the fate of all disinterested pursuits. What he sought to uphold was not this or that discipline, but the kind of discipline which, suitably directed, enabled man to have a mind at all. Man, he declares, 'must unify his life somehow', and this can be done 'either positively, by bringing every activity into harmony, or negatively, by suppressing all but one'.[9]

The book purports to outline a Map of Knowledge. Some critics have assumed that in isolating five forms of knowledge — art, religion, science, history and philosophy — Collingwood was attempting to cut up knowledge into slices. To do so would be to credit each with a particular 'slant' on reality. This is to misunderstand his purpose. Collingwood is at pains to demolish 'the notion of a map of knowledge distinct from knowledge itself'. Except in abstract thought, the five forms of knowledge are not separable; they are 'moments' in a rhythm. To possess true knowledge is to engage at once in art, religion, science, history and philosophy: for in pursuing these activities, we are achieving progressive self-awareness. 'In the

[8] *Speculum Mentis*, p. 108n. [9] Op. cit., Prologue, p. 37.

toil of art, the agony of religion and the relentless labour of science, actual truth is being won and the mind is coming to its true stature.'[10] The process is one of education. What, we may ask, is the end of education? It is certainly not to know quantities of science, history, etc.: which is merely to stoke the engine. On the contrary, the end of education is the achievement of a state, a condition. And this state or condition, if we can imagine it at all, bears a much closer resemblance to ignorance than to knowledge; for 'at the completion of its education, a mind would step forth as naked as a new-born babe, knowing nothing but having acquired the mastery over its own weaknesses, its own desires, its own ignorance, and able therefore to face any danger unarmed'.[11]

The language of the above passage, like that of *Speculum Mentis* as a whole, has perhaps lost some of its appeal. One reason is our suspicion that in philosophy well or brilliantly written we are being wooed rather than convinced. That 'idealist' philosophy is better composed than 'realist' has often been suggested. With the signal exception of Hume, the idealists have perhaps been the better stylists, which does not mean better writers of philosophy. In a fine passage in the *Essay of Philosophical Method*, Collingwood contrasts the work of the poet and that of the philosophical writer. The one he describes as engaged in producing a jewel to be admired, the other a crystal in whose depth the thought can be seen without flaw or distortion. In short, while the poet's trade is that of jeweller, the philosopher's is that of lens-grinder.[12] Although deliberately understressed and 'flat' writing is the fashion among certain critical or analytical thinkers, the result is not necessarily a more rational or impartial approach: the emotion is there, influencing however slightly the argument and winning the reader to a point of view. There is as much propaganda in Wittgenstein's *Tractatus Logic-Philosophicus* as in the *Summa Contra Gentiles*.[13] Between style and thought there exists an evident correlation; a profound thinker must perforce write in depth. A superficial thinker, however graceful his writing appears always *outside* his subject. These observations are made to point a moral in Collingwood's case upon which we have already touched: Collingwood is not so much writing about philosophy, as engendering by his writing the philosophical mood or attitude. Words, as a Hindu thinker has said, are employed to get beyond words; but only the mastery of words, the shaping and domination of language, will permit their supersession. *Speculum*

[10] Op. cit., p. 312. [11] Op. cit., p. 316.
[12] *Essay on Philosophical Method*, p. 213.
[13] Collingwood makes this point very effectively in another connection in *The Principles of Art*, p. 264.

Mentis is a passionate denunciation of philosophy as an academic study in favour of philosophy as a way of life. Thinking must pass into action in order to become thinking again; and thinking which is of the whole man is already action.

III

In the interval between the publication of *Speculum Mentis* and the *Essay on Philosophical Method*, Collingwood was occupied with teaching, reflection and much archaeological research. All were undertaken with his customary thoroughness: *engagement* was characteristic of all his labours and even of his pastimes. Having commented upon the impact of the *Essay*, we do not intend to linger over it in detail. In the chapter on 'Philosophy as Categorical Thinking', however, Collingwood embarks upon certain general considerations of extreme interest. Declaring that logic is committed to the Ontological Proof, since it is 'about itself', he advances two contentions: first, that thought, 'when it follows its own bent most completely and sets itself the task of thinking out the idea of an object which shall completely satisfy the demands of reason', is 'never devoid of objective and ontological reference'; and secondly, and as a consequence of this, that 'essence involves existence, not always, but in one special case, the case of God in the metaphysical sense: the *Deus sive Natura* of Spinoza, the Good of Plato, the Being of Aristotle: the object of metaphysical thought'. From the combination of these two points he arrives at the following conclusion: 'All philosophical thought is of the same kind, and every philosophical science partakes of the nature of metaphysics, which is not a separate philosophical science but a special study of the existential aspect of that same subject-matter, whose aspect as truth is studied by logic and its aspect of good by ethics'.[14]

This passage, especially its remarks upon logic, has received some rough handling.[15] Its interest lies not so much in its originality (if indeed it is original) as in its later repudiation by its author. The *Essay on Metaphysics*, written seven years later, summarily demolishes Aristotle's idea of Being. Along with it is traduced the 'object of metaphysical thought' as here defined. Yet we know that as late as 1936 (and the notes for the *Essay on Metaphysics* date from 1938), Collingwood still apparently regarded metaphysics in the traditional sense, that is to say, as a science concerned with 'the One, the True, the Good'.[16] Whether even after completing such books as the *Essay*

[14] Op. cit., pp. 127, 131.
[15] e.g. 'Mr Collingwood and the Ontological Argument', G. Ryle, *Mind* (1935).
[16] T. A. Knox, Editor's Preface to *The Idea of History*, p. x.

on Metaphysics and the *Autobiography*, he would have returned to the subject, as one that fascinated him, we never can know. In view of his total repudiation of the traditional view, it is difficult to see how, short of confessing to have strayed, he could have done so. It may be that in writing the *Autobiography*, a work obviously composed in some haste, he was impelled by two motives. Overtly he wished to put on record his considered views on history; he admits as much.[17] But if so, we are led to wonder, as I have already mentioned, why he did not sit down to finish *The Idea of History* itself. Much of that work, originally to be called *The Principles of History*, was already written. The real motive, though concealed, may have been the following: he was dissatisfied with the direction his thought had latterly taken. He wished to institute an inquiry into his own philosophic record. Obsessed with history, and having defined all history as the history of thought, he was anxious to lay bare the history of his own thought. This, if true, would account for another significant fact. In the course of the investigation, the record suffers perceptible distortion. We are led to believe that the theory of presuppositions as expounded in the *Essay on Metaphysics* was a theory to which he had always subscribed. In short, the *Autobiography* endeavours to demonstrate a measure of consistency in argument which, at the time of writing, had already undergone violation.

To suggest that Collingwood's major error was that of historicism is to imply that he had fallen into an error against which he had issued a warning not merely in the *Essay on Philosophical Method* but in *Speculum Mentis*. The warning is one which we have quoted. The 'unity of knowledge' is sought not positively, 'by bringing every activity into harmony', but negatively 'by suppressing all but one'. Most critics have been inclined to leave the matter at that. In the remainder of this essay, however, I propose to raise two questions. First, I wish to ask whether the matter can be left at that. Secondly, I wish to suggest that the wrong turning taken by Collingwood was due less to his concentration upon history than to his revulsion from natural science, or rather from the method associated with natural science.

Granted that Collingwood's theory of history and historiography was distorted, what of his practice of these things? Here we seem about to take leave of philosophy. Admittedly, the subject is one to which none but the professional historian is entitled to give an answer; and so few persons enjoy Collingwood's equal competence in philosophy and history that the subject has received scant attention. Yet following upon Collingwood's death, a former colleague,

[17] *Autobiography*, Chapter x.

whose views are entitled to respect, pronounced upon this very subject: I refer to the highly critical, if generous, appraisal by I. A. Richmond entitled 'Appreciation of Collingwood as an archaeologist' (*Proceedings of the British Academy*, vol. XXIX). From this account, we realize that Collingwood's view of history was not merely productive of theoretical error but involved serious errors in practice. According to Richmond, some of Collingwood's mistakes were due to rigorous application of arbitrary philosophical axioms: the simplified 'we find what we are looking for' became at a more sophisticated level the logic of question and answer, and finally the theory of presuppositions. In fact, theory affected practice and practice affected theory: to that extent Collingwood would have approved. Another curious inconsistency, though one to which Richmond does not refer, may be observed in the volume in the *Oxford History of England* on Roman Britain, of which Collingwood wrote the first four books.[18] The first chapter is devoted to an inquiry into the influence of geographical conditions upon the inhabitants of Britain prior to the Roman invasion. Collingwood attributes certain aspects of the British genius, such as its combination of conservatism and initiative, to the fact that Britain is partly in the 'highland zone'.[19] In the *Idea of History*, on the other hand, he takes pains to demonstrate that geographical (and therefore physiographical) conditions never operate as a 'cause'. Which of these contentions is correct does not matter: the point is the obvious inconsistency.

Such a criticism of Collingwood as a historian — which, since he came to identify philosophy and history, implies criticism of his philosophy — must not be taken to suggest that his works on Roman Britain were failures. The little volume entitled *Roman Britain* (1923), of which he undertook three revisions in nine years, is a minor classic. Moreover, no student can ignore his fine contribution to Frank's *Economic Survey of Rome* or the volume in the *Oxford History of England* already mentioned. What is of primary interest, given his aberrations of *theory*, is the fact that he writes history much as other historians do. The strictures of Richmond apply primarily to his archaeological work. This is significant: for it is in the sphere of archaeology that the logic of question and answer and the theory of presuppositions operate with least effect. Arnold Toynbee's comments on Collingwood's method are not so telling.[20] Seizing upon Collingwood's statement that history is the 're-enactment' by the historian of a past

[18] *Roman Britain and the English Settlements*, by R. G. Collingwood and J. N. L. Myres (1936).
[19] Op. cit., pp. 3-4.
[20] *A Study of History*, vol. IX, Annexe, pp. 718-37.

event, Toynbee takes the idea of re-enactment literally. Having done so, he proceeds to suggest that a historian endeavouring to understand Tamerlaine's slaughter of 5000 persons at Zirih for the purpose of using their skulls to build minarets is obliged, on Collingwood's theory, personally to undertake a similar act of multi-homicide: whereas Collingwood repeatedly states that the re-enactment must take place in the historian's own mind, and that the 'historical imagination' is precisely that faculty which, in the best historians, permits this to happen.[21] The whole of the Annexe would seem to betray a misunderstanding of Collingwood's central thesis. Furthermore, Toynbee's assertion that Collingwood's ideal form of knowledge was mathematical knowledge is at variance with the whole tenor of his teaching. On the other hand, Toynbee is surely right to draw attention to one particular error of Collingwood's, namely his insistence that the historian can re-enact only past *thought*. Feeling and will are also grist to the historian's mill: indeed, much of Collingwood's own historical writing, such as the *Oxford History* volume, is concerned with epochs about which we know so little of individual personalities that there are few discernible 'acts done on purpose' (another of Collingwood's definitions of the subject-matter of history) to study and assess.

IV

On the danger of introducing morals into history much has been written. There are equal dangers, however, in their total exclusion. Collingwood's historicism is an example of such exclusion pushed to the point at which the past is deprived of significance. In the chapter on Roman Britain in the *Autobiography* (xi), he lays down his rules for the archaeologist. 'Whenever you find any object', he writes, 'you must ask, What is it for? and, arising out of that question, Was it good or bad for it?, i.e. Was the purpose embodied in it successfully embodied in it, or unsuccessfully?' This is advanced as being the whole duty of the historian. Now if metaphysics is history, it must follow that all metaphysical questions must be treated in a similar manner. Thus, to ask such questions as What is Duty? or What is the Good? is strictly to ask not a general question but a particular one: e.g. What did Kant think was Duty? or What did Plato think was the Good? And this is precisely what Collingwood, repudiating his definition of the purpose of metaphysics in the *Essay on Philosophical Method*, maintained. Questions such as What is Knowledge?, What is the *summum bonum*? or What is Art?, he says a few pages

[21] Cf. *The Idea of History*, pp. 326-7.

earlier in the same chapter, are pseudo-questions. It is here surely that historicism breaks down under its own weight. For to examine the Kantian view of Duty is not merely to ascertain what Kant actually said about Duty (which would be to refer to a series of printed passages) but to take into account the fact that Kant, in troubling to say it, believed that he was saying something true. (We recall Collingwood's earlier statement that thought, pursued to its goal, has 'ontological significance'.) Kant, like Plato or indeed Collingwood himself, was not simply saying what he happened to say: nor was he saying what the historical context obliged him to say: he was saying what, on the basis of reflection upon the nature of the 'subject-matter' before him, he felt he ought to say. If, for example, we apply Collingwood's questions to such a problem as What is Purpose? (a question frequently asked in philosophical biology), we have first to ask the question What did Darwin (for instance) think was Purpose? and secondly Was the purpose of asking What is Purpose? successfully achieved? Now if the general question What is Purpose? is a pseudo-question, the particular question What did Darwin think was Purpose? is equally a pseudo-question. It will not do to argue that Darwin should not have asked so broad a question; for even if Darwin's question were dismissed as a pseudo-question, the historian, in Collingwood's view, must still ask what purpose he had in mind in raising it. And unless he knows or has some idea what Purpose is, he will be unable to understand the nature of human motive. The same applies to the concept of Duty. In saying what he felt he ought to say about Duty, Kant felt he had a duty to say it; and unless we can ask the general question What is Duty?, we cannot on any one occasion distinguish a man's considered conclusions from the errancy of caprice. Finally, if such general questions as What is Metaphysics? or What is Art? are declared to be pseudo-questions, Collingwood's *Essay on Metaphysics* and his *Principles of Art* must be concerned with pseudo-questions.

There is some evidence to suggest that Collingwood, even at the time of expounding his later theory of metaphysics, realized what was happening. In the chapter on 'Theory and Practice' in the *Autobiography*, he defines an action which procures 'ends beyond itself' as an economic action. Now the double question which Collingwood enjoins upon the historian is applicable only to economic actions: for the question What is it for? and Did it successfully embody what it is for? are questions relating to the end of which an object is a means. In short, Collingwood is using good and bad in this context in a utilitarian sense. But when such questions are applied to philosophical theories, such as the Kantian theory of Duty or the Platonic

summum bonum, the utilitarian approach fails to yield illuminating results. If we ask in respect of Kant or Plato what their theories were for, we are obliged to answer that they were for the enlightenment of mankind as to the nature of moral action; but if we cannot ask any question of an action except in terms of means and ends, we are offering, in Collingwood's own words, 'an account of economicity when professing to offer one of morality'.

In fact Collingwood provided a refutation of this ethical historicism (which is another name for utilitarianism) in an unexpected place. In his lively account of a cruise among the Aegean islands, *The First Mate's Log* (1940), he examines this very question. The question arises in the course of an argument as to the merits of the monastic life. Defending that life for certain people, Collingwood argues that some forms of life are not to be assessed by asking what they are for or what purpose they serve: they are good in themselves. And he thereupon proceeds to a most effective polemic against the utilitarian view of action. Such a view, he maintains, implies that everything has an exchange-value: but if that were so, there would in the end be no point in exchanging anything. It would be like having a paper-currency which no one could cash. Now the same argument, it is obvious, must apply to his theory of metaphysics: for if all theories, whether of science or morals, are relative to the epoch in which they originate, their value is purely utilitarian or pragmatic, and this would apply to the theory of metaphysics which pronounces them to be so, since this theory is itself the product of a particular epoch with its particular constellation of presuppositions. At some point the paper-currency of theories must be 'cashed': we must ask which is the *right* view. And if the answer is 'Collingwood's theory of presuppositions', then it is no longer true to say that all absolute presuppositions are relative: indeed the expression itself is contradictory. If the past is relative, then it has no significance, and the historian-philosopher who deprives the past of significance is hardly deserving the name of historian or philosopher.

V

Collingwood's writings on history are characterized by attacks on the subservience of history (and likewise of philosophy) to natural science. In the concluding years of his life, these attacks on natural science became more frequent and bitter. But Collingwood was not an enemy of natural science. On the contrary, he was a keen student of it. What he was attacking was the *method* according to which natural science had been pursued. And his attack was directed not

against the results of this pursuit, for these had hitherto proved remarkable; his attack was directed at those practitioners who, encouraged by these remarkable results, believed that the method was universally applicable. The result, among other things, was what he called 'scissors and paste' history. In 1934 and again in 1937 Collingwood had delivered courses of lectures at Oxford on 'Nature and Mind'. These lectures, as those who heard them can testify, were of exceptional brilliance; they were possibly more revelatory to the student than those on the 'Philosophy of History'. When the question arose as to their posthumous publication, the Editor found little that needed change; the printed version differed from the spoken only in so far as the concluding section, that occupied with Collingwood's own cosmological speculations, was omitted. Dissatisfied with his original statement, Collingwood substituted a shorter section concerned with the transition from Nature to History.

The transition from Nature to History represented a transition from a science approached from the physico-mathematical point of view – the 'method', Cartesian in origin, to which Collingwood referred – to a science where this point of view was seen to be irrelevant and indeed pernicious. Now Collingwood shows himself satisfied that the scientist, in discussing Nature, should employ the physico-mathematical method. He even goes so far as to argue, in Chapter x of the *Autobiography*, that nineteenth-century natural philosophers were right to adopt this *mechanistic* approach. 'Natural science as it exists today,' he writes, 'does not include the idea of purpose among his working categories.' Yet those sections of *The Idea of Nature* which deal with Alexander and Whitehead discuss with approval the ideas of *nisus* and 'lures': terms which, unless they are to be regarded as pseudo-concepts, embody the idea of purpose, and indeed suggest that all processes within nature are purposive. Moreover, in the cosmological speculations with which he originally concluded his lectures, he advanced a theory according to which Nature was engaged, by the progressive overcoming of 'externality', in turning into Mind. Unless, therefore, the transition from one to the other was characterized at some point by a complete break, the world of Nature must already be shot through with purpose. In either case, any study of Nature which followed a method incompatible with the idea of existence of purpose was doomed to failure.

That mechanistic processes are found in nature (e.g. in the reflexes of physiology) is generally admitted: but such mechanistic processes are as it were 'contained' by organic processes. Moreover, it will be found that all attempts to analyse or reduce these organic

processes to terms of mechanism break down. Had Collingwood pursued his study of cosmology, he would no doubt have realized that the foundation of Nature was not History ('Where do we go from here? he asks at the end of *The Idea of Nature*, and answers: 'We go from the idea of Nature to the idea of History'), but rather that the foundation of History was a world of organic nature. For if history is the 'self-knowledge of mind', this mind has developed self-knowledge only by having issued from an organic process itself characterized by the idea of purpose, which at the stage of mind has become a conscious purpose. By going from Nature to History, Collingwood appears to mean only that historians are able to ascertain that at certain times and for certain reasons scientists have propounded certain theories of nature. This is simply his application of the theory of metaphysics as the science of absolute presuppositions historically ascertainable. But a mind that 'knows itself' only by knowing that other people's minds have entertained certain 'absolute presuppositions' which are neither true nor false is not a mind that is able to 'grow' by such knowledge. Our minds can grow only by the acquisition of truth, or they are no more than Lockean blank sheets upon which certain 'data' are printed. What is printed must be read, understood and judged; and there is no judgment without the application of criteria of truth and falsehood. It is a melancholy reflection that as early as *Speculum Mentis* Collingwood had denounced the point of view according to which the 'Map of Knowledge' is conditioned by 'a group of presuppositions [*sic*] which lie beyond and behind its very beginning'.[22]

The lesson that Collingwood fails to bring out in *The Idea of Nature* is the lesson that not merely history but natural science itself has begun to suffer from the exclusive application of the physico-mathematical method. And owing to his theory whereby the philosopher was able merely to ascertain what particular theory of science had been held at a particular time — 'interpreting it, subject to certain rules, in such a way as to satisfy himself that the man whose work it records really did observe what he professes to have observed' — there is no occasion apparently to ask whether or not it is true. 'Metaphysics . . .' Collingwood summed up in a well-known passage, 'is no futile attempt at knowing what lies beyond the limits of experience, but is primarily at any given time an attempt to discover what the people of that time believe about the world's general nature, such beliefs being the presuppositions of all their "physics", that is their inquiries into its detail.' One notes the word 'primarily'; one notes again the absence of any suggestion that, in addition to

[22] Op. cit., p. 309.

discovering what people thought, the metaphysician should ask whether such theories were true. Indeed, we know in advance that truth does not enter into the inquiry, because the presuppositions are by definition neither true nor false.

By way of conclusion I propose to raise a matter which is naturally a subject of speculation: my views as here stated must therefore remain tentative. I have suggested that Collingwood failed to complete *The Idea of History* because he lacked the 'metaphysics' to do so. I am now going to suggest that he jettisoned the conclusion to *The Idea of Nature* because, in working it out, he found himself tempted to return to the metaphysical position which he had left behind in the *Essay on Philosophical Method.* This, I believe, may have led to the dissatisfaction of which I have spoken. To quote from notes in my possession would be improper: Collingwood made it clear, and his executors have confirmed, that such writings as he did not expressly designate for publication should not be allowed to see the light. All I may legitimately affirm is that the sketch of a cosmology which he outlined in his lectures seems to imply a view of the mind capable not merely of elaborating a 'Map of Knowledge' but of attaining to knowledge itself — of apprehending 'the One, the True, the Real'. And I would further suggest that had he lived, the metaphysical historicism or historicist metaphysics into which he had become involved in consequence of his immersion in historical studies, would have been gradually abandoned. For the study of Nature and Mind is a study of the Manifold and the One, and no such study could reach maturity without a return to the classical metaphysics of Aristotle, with its culmination in the idea of mind as identified with its object. 'The intellect is truer than science', Aristotle declares in the *Metaphysics*, for to Aristotle metaphysics was not the science of presuppositions but the science of intellectual intuition.

To make comparisons between Collingwood and his contemporaries is a difficult task. He has been compared with Dilthey and with Husserl, even with Carnap. The philosopher to whom he bears greatest resemblance is, in my view, none of these. To me he bears a close resemblance to that great and versatile man, Ernst Cassirer. With Cassirer, as with Collingwood, one follows with mounting admiration and respect the course of a masterly argument, ballasted with learning and enlivened by wit. Yet one feels, at the end of every work, that the last word has not been said. Cassirer's monumental work *The Philosophy of Symbolic Forms* has its counterpart in *Speculum Mentis. The Myth of the State* has its counterpart in *The New Leviathan. Sprache und Mythos* has its counterpart in *The Principles*

of Art. *The Problem of Knowledge* has its counterparts in *The Idea of Nature*, the *Essay on Metaphysics* and *The Idea of History*. *The Problem of Knowledge* remained unfinished, just as in a sense did the three books of Collingwood just mentioned. Even Cassirer's *Essay on Man*, a kind of later digest of the *Philosophy of Symbolic Forms*, is an unfinished work, because the view that man is a 'symbolic animal' no more satisfies than the view that man is an 'historical' animal. At some point the code must be broken. Symbol and historical presuppositions must give way to truth and fact. The last word is not said . . . But then it never is.

THE MARXIST-LENINIST THEORY OF RELIGION

By H. B. Acton

I

Lenin expounds what he takes to be the Marxist view of religion in an article entitled *Socialism and Religion*, first published in 1905, and in an article entitled *The Attitude of the Workers' Party towards Religion*, first published in 1909. But of course he refers to the subject in many other writings, such as his *Materialism and Empirio-Criticism* (1909) and *Philosophical Notebooks* (various dates about 1914). Lenin also discussed the subject with Gorky in a correspondence between them in 1913. One is inclined to say that Lenin was obsessed with an anti-religious fury which overflows into his writings at the slightest encouragement.

In *Socialism and Religion* Lenin writes that religion is 'one of the forms of spiritual oppression that everywhere weighs on the masses of the people . . .' He goes on to say that belief in a better life beyond the grave inevitably arises when the proletarians fail in their struggle against their exploiters, just as belief in God, devils and miracles arose when savages failed in their struggle with nature. Furthermore, 'to those who live on the labour of others religion teaches charity on earth, offering them a very cheap justification for their whole existence as exploiters and selling them at a suitable price tickets for admission to heavenly bliss. Religion is the opium of the people. Religion is a kind of spiritual gin in which the slaves of capital drown their human shape and their claims to any decent human life.' In *The Attitude of the Workers' Party towards Religion* he writes as follows: 'The deepest root of religion today is the social oppression of the working masses and their apparently complete helplessness in the face of the blind forces of capitalism . . . "Fear created the gods". Fear of the blind forces of capital – blind because it cannot be foreseen by the masses of the people – a force which at every step in life threatens to inflict and does inflict on the proletarian and small owner "sudden", "unexpected", "accidental" destruction, ruin, pauperism, prostitution and death from starvation – such is the *root* of modern religion which the materialist must bear in mind first and foremost if he does not want to remain an infant-school materialist . . .'

Having stated that this is what religion is and how it arises, Lenin goes on to consider what should be the attitude of the 'workers' party' towards it. In brief, his view is that it is pointless to take part in direct attacks on religious beliefs, since these beliefs inevitably

arise in certain sets of social circumstances, and will cease to be entertained only when these circumstances have been transformed and a classless society has been instituted. He urges his readers not to 'fall into the abstract and idealist error of arguing the religious question from the standpoint of "reason", apart from the class struggle – as is not infrequently done by bourgeois radical democrats'. While, therefore, those who direct Party affairs will, as Marxists, always be atheists, they should not antagonize possible supporters by atheist propaganda, but should even welcome into the Party any priests who are willing to support its social aims – as long as these priests do not use their position to gain religious converts from within the ranks of the Party.

Marxists may write books and pamphlets criticizing religion, but such things must be subordinated to the basic task of the Party – 'the development of the class struggle of the exploited *masses* against the exploiters'. Reactionary social groups may even welcome atheist propaganda on the part of Marxists, in the belief that it could cause divisions among the workers. The leaders of the 'workers' party' should, in Lenin's view, consider the situation concretely, dialectically. He explains that the statement in the Erfurt Programme of 1891 that religion was a 'private' matter, means that a Socialist government would not prohibit religious practices (for did not Bismarck actually *stimulate* Roman Catholicism by attacking it?), but not that an instructed Marxist could possibly be a religious man. 'Marxism', he writes, 'has always regarded all modern religions and churches and all religious organizations as instruments of bourgeois reaction that serve to defend exploitation and to drug the working classes.'

Nor is Lenin at all favourable towards the idea of improving religion, or of inventing a new one. In an article entitled *On the Significance of Militant Materialism*, which appeared in the Russian periodical *Under the Banner of Marxism* in 1922, Lenin attacks Arthur Drews, who in his book, *The Christ Myth*, had urged the need for a regenerated religion freed from discredited mythology. 'This', writes Lenin, 'is a straightforward conscious reactionary who is openly helping to substitute new, viler and more despicable religious prejudices for old and rotten ones.' Lenin's views on some ideas for a new and purified religion are forcibly expressed in some letters he wrote to Gorky in 1913. Gorky, it appears from what Lenin quotes, had defined God as a complex of ideas which 'arouse and organize social sentiments with the purpose of binding the individual to society and of bridling animal individualism'. But Lenin will have nothing of such talk, which he considers 'reactionary' because it

resurrects 'the clerical-servile idea of "curbing animality" '. 'In fact', he says, 'the idea of God has always blunted the social emotions by substituting concern for the dead for interest in the living. It has *always* involved the idea of slavery (of the worst and most hopeless slavery). The idea of God has never "united the individual with society". It has always bound the oppressed *classes* by faith in the divinity to submission to their oppressors.'

The views thus attacked went at the time under the name of 'God-building'[1] and had been publicly supported by the Russian Social Democrat Lunacharsky, who had written approvingly of 'the deification of the higher human potentialities', of 'religious atheism', of 'scientific socialism in its religious significance'. In *Materialism and Empirio-Criticism* (in the Preface) Lunacharsky was stigmatized as one of 'the would-be Marxists who have . . . undertaken a veritable campaign against the philosophy of Marxism, and in Ch. VI, section 4, Lenin asserts that, however good Lunacharsky's intentions may be, his new religion must in fact serve the class interests of the bourgeoisie by bolstering up the religious attitudes which support them. The main theme of this book, of course, is that philosophical idealism is a form of religious apologetics, and that a materialist philosophy is the only one which is at the same time true and socially advantageous for the mass of mankind.

This view is repeated over and over again in Lenin's *Philosophical Notebooks*. It is significant that among the works there commented upon is Feuerbach's *Vorlesungen über das Wesen der Religion* (Lectures on the Essence of Religion) (1851).[2] Alongside Feuerbach's sentence: 'Hence Christians tear the spirit, the soul of man out from his body and make this sundered, decorporealized spirit their God' Lenin writes: 'Decorporealized spirit = God.' He also copies out and emphasizes the following interesting passage: '. . . to turn friends of God into friends of mankind, believers into thinkers, worshippers into workers, candidates for the other world into students of this world, Christians who, in accordance with their creed and confession, are *half animal, half angel*, into *men*, into *complete* men . . .'

[1] For a brief account of 'God-building' and (a different view) of 'God-seekers', see G. A. Wetter, *Der Dialektische Materialismus. Seine Geschichte und sein System in der Sovjetunion* (Freiburg, 1952), pp. 100-2. Lenin refused to distinguish between these different views or between them and orthodox religions. They differed, he said, only as blue devils differ from red devils.

[2] Marx's views on religion had, of course, been determined by Feuerbach's earlier writings, especially by *Das Wesen des Christentums* (The Essence of Christianity) (1841), *Vorläufige Thesen zur Reform der Philosophie* (Preliminary Theses towards the Reform of Philosophy) (1842), and *Grundsätze der Philosophie der Zukunft* (Foundations of the Philosophy of the Future) (1843).

II

We may start our discussion of these views by considering more closely the gin and opium *motif.* According to a preliminary, superficial interpretation of this part of the Marxist theory, the bourgeoisie uphold religious beliefs and practices in order that the proletarians whom they exploit shall put up with their sad plight on earth in the hope of getting release and compensation in heaven. On this view, the opium and gin are *administered by* the exploiters *to* the exploited classes so that they may relieve their sufferings in compensatory dreams instead of by means of revolution. Priests administer religious anaesthetics under instructions from their capitalist employers. But Lenin was too good a Marxist to believe this, at any rate in the crude form in which I have expressed it. The phrase 'Religion is the opium of the people' was used by Marx in an article entitled 'Zur Kritik der Hegelschen Rechtsphilosophie. Einleitung' (Criticism of Hegel's Philosophy of Right. Introduction) which appeared in the *Deutsch-Französische Jahrbücher* in Paris in 1844. It is important to notice, however, the whole passage in which it occurs, which reads as follows: 'Religion is the sigh of the hard-pressed creature, the heart of a heartless world, the soul of soulless circumstances. It is the opium of the people.'[3] From this it appears that the opium is not administered to the people, but rather that they resort to it themselves in order to relieve sufferings for which they know no other remedy. Thus according to Lenin those who suffer the evil effects of capitalism and do not know how to prevent them find imaginary remedies in a world of fantasy where the humble and meek are finally exalted. It will be objected that it is not only in capitalist society that religious beliefs have been held. In pre-capitalist society, however, they were, according to Lenin, the result of man's ignorance of the laws of the physical world. In early society religions arose from fear of natural catastrophes and ignorance of how to deal with them. But by the nineteenth century natural science had developed to such an extent that in western Europe and America this source of religious belief was being eliminated, and it was now ignorance and fear of *social* catastrophes which gave rise to it. Furthermore, the bourgeoisie were helping to foster religious beliefs in the very process of exploiting the proletarians. For whether or not they themselves share these beliefs, by their very existence as profit-making and therefore (according to Marx) exploiting enterprisers they unwittingly produce

[3] *Marx-Engels Gesamtausgabe* (*MEGA*), I, I, I, p. 607. When he wrote this Marx may well have had in mind the following sentence by Sebastian Franck quoted by Feuerbach in Ch. XIII of *The Essence of Christianity*: 'God is an inexpressible sight deep down in the souls [of men].'

the conditions in which bewildered and distressed proletarians turn to religion for solace and compensation.

III

Let us now examine the more sophisticated version of the opium theory, the version that Lenin inherited from Marx.

(*a*) It is a consequence of this view that religion will decay when the social conditions which force people to have recourse to it have ceased to exist. To attack religion without also attacking the social order within which it grows is, on this view, to substitute ideological shadow-fighting for genuine social revolution.[4] Lenin makes it quite clear, therefore, that the prime concern of members of the Communist Party is to overthrow capitalism and to establish socialism. This being so, it is a matter of expediency (this is what Lenin means — here, as often elsewhere — by considering the situation 'concretely, dialectically') whether or not, on particular occasions, religious ideas should be opposed, ignored or even, here or there, given some temporary encouragement. Clearly, when the exploited proletariat have deep religious feelings (as they are *bound* to have, according to this theory) then Marxist propagandists may antagonize those they wish to organize if they stress their own radical opposition to the religious view of things.

(*b*) There are serious difficulties in the view that religion would spontaneously die out once the capitalist system has been replaced by a socialist or communist order. The basis of the whole Marxist theory is that religious illusions will inevitably impose themselves when individuals can see no other way out of their real misfortunes. Before capitalist exploitation existed men were religious, according to the Marxists, from fear of natural disasters. They can therefore only avoid the need for some other form of religion when capitalism has been overcome if in the communist society there are no disasters, nothing to fear, no fundamental needs that remain unsatisfied. Communism, of course, is defined by Marxists as a form of society in which each member receives according to his needs and contributes according to his abilities. That is to say, it is a form of society in which everyone does what is expected of him and gets what he wants.

Definitions, however, do not establish real existence in the future any more than they do in the present or past, and it is reasonable to suppose that in any form of society it is worth our while to consider

[4] Marx, *MEGA*, I, I, I, p. 608. 'Criticism of religion is therefore, in germ, a criticism of the vale of tears whose halo is religion.'

men will suffer severe disappointments, and will grieve when their friends withdraw their friendship or when they die. In any form of society it is worth our while to consider men will not want to suffer from disease and will, for the most part, prefer living to dying. Scientific and social discoveries have, of course, lessened the amount of disease and delayed the onset of death. But death is there in any event, and dying none the better, in many cases, from being so long delayed. It may be argued that I am projecting into a future communist society the *individual* fears that belong to bourgeois society, and that in a communist society men would cease to fear their own dissolution because they would have identified themselves with mankind. But I suggest that this sort of argument loses sight of the close connection of each person with his own body, and suggests a moral absorption in mankind that could only occur if mankind had a single body.

In any case death is not the only source of deep anxiety. Many people are unhappy when they have lost, or have failed to obtain, the love or regard of others, and it is hard to believe that any social order will guarantee that they will get it. Of course the Marxist contention is that once disease is greatly diminished and poverty and exploitation abolished, hatred, aggression, envy and frustration, and the sense of individual failure and personal worthlessness will no longer have the soil in which to flourish. In particular, it is the assumption of the Marxist view that with the abolition of economic exploitation all exploitation will cease, so that there need be no fear, for example, that the more intelligent will make use of their greater ability to exploit the less intelligent. I can only say that, were this nevertheless to happen, then the position of the less intelligent would be hopeless indeed, since there would be nothing they could do, short of destroying the social order by brute force, to counter the manœuvres of their cleverer associates. The Marxist answer to this might be that where everyone has a sufficiency, and where no one, therefore, suffers from the physical effects of poverty, exploitation can no longer be said to exist, and hence the merely *relative* sufferings of the not-so-well-off would not be severe enough to give rise to religious hopes and fantasies. But we are then brought back to the view that the primary basis of religious beliefs and practices is the fear of physical suffering and death and the hope of obtaining supernatural alleviation. Yet physical suffering and death are surely inevitable features of any society of embodied creatures. It may seem strange to accuse Marxists of taking too *etherial* a view of human possibilities, yet that is just what they are doing when they say that in communist society there will be no need for religious hopes.

This tendency to believe that men may some day free themselves from their biological limitations shows itself in a striking passage in the *Anti-Dühring* where Engels says that in the future communist society 'the struggle for individual existence comes to an end. And at this point, in a certain sense, man finally cuts himself off from the animal world, leaves the conditions of animal existence behind him and enters conditions which are really human' (Penultimate paragraph of Part III, Chapter II).

(*c*) We must now give some attention to the mechanics of the process by which religion is held to be the outcome of ignorance and fear. The views of Marx and Engels about this were greatly influenced, as is well known, by those of Feuerbach. Feuerbach believed that there is one world only, the natural world, and that we are all, therefore, in it and parts of it.[5] He held, furthermore, that men's concern with a supernatural world was morally harmful as well as intellectually misleading, for their concern with a non-existent world distracted their attention from the tasks that it was their duty to face in the real world.[6] The belief that the soul survives death in a heaven where justice is secured arose from the prevalence of injustice on earth, and was hence an imaginary compensation for real sorrows.[7] The idea of God was that of a being with all the powers that men would like to have. Hence, in worshipping God men were really celebrating human achievement and progress, although they thought they were addressing the supernatural creator of the universe.[8]

It is important for us to notice the following features of this view. One point is that although really men created God, they believe that God created them. They are, in a sense the slaves of their own

[5] 'The theoretical basis, the subject of the new philosophy is, therefore, not the "I", not absolute, abstract spirit, not Reason *in abstracto*, but the actual and whole being of man . . . It is a man who thinks, not the "I", not Reason.' *Foundations of the Philosophy of the Future*, 50.

[6] In *The Essence of Christianity*, however, Feuerbach does not call for the abolition of religion. Indeed, he was criticized by Stirner and others for appearing to favour a new, atheistical religion. This aspect of his work can be seen in the very last pages of *The Essence of Christianity* where, in connection with the Eucharist, he says that eating and drinking *ought to be* religious acts. In the Preface to the same book he says that he has made an analysis of religion that is quite unpolitical. He believed that when mankind had grasped the nature of *unselfconscious* religious beliefs and practices they would be in a position to deal religiously with human relationships. As H. Arvon wrote in *Ludwig Feuerbach ou la Transformation du Sacré* (Paris, 1957), p. 131: 'Feuerbach avait voulu transformer le sacré, Marx le chasse.'

[7] 'Belief in heaven, or, more generally, in a life beyond, rests on a *judgment*. It expresses both praise and blame; it is critical in nature; it puts together a bouquet out of the flowers of this world . . . This other life is nothing else than life in agreement with feeling, with the Ideal with which this life is in conflict.' *The Essence of Christianity*, Chapter XIX.

[8] 'Man — this is the essence of religion — objectifies his own nature, and then makes himself an object of this objectified nature transformed into a subject and a person.' *The Essence of Christianity*, Chapter II.

creations, haunted by their own handiwork. Another point is that religious beliefs, theology and metaphysics are regarded as works of the imagination, elaborated, of course, by rational arguments, but operating in the service of wishes.[9] But these things, it is thought, become apparent when human beings are studied as natural beings, and hence belief in a supernatural world will disappear when a science of man (called 'anthropology' by Feuerbach) has been established. Feuerbach also thought that the belief that man is both a soul and a body is the result of his misunderstanding of his own nature and would cease to be held when he had achieved self-knowledge.

We can now see the sort of mechanism by which religious belief is held to be established. In the imagination one can travel in an instant from London to Jerusalem (Old or New), and for this sort of travel there is no need to consider the means.[10] But if we wish *really* to travel very fast, then we must invent, test and improve certain machines. According to the Marxists, the religious man is like someone who satisfies himself with imaginary journeys on magic carpets instead of putting himself to the trouble of inventing and operating effective material means. He lives, at any rate from time to time, in a paradise of his own making. He takes opium believing it to be the elixir of life, but if he had had enough to eat he would not have concerned himself with either. An insubstantial dream world is called into existence in order to restore the balance of reality. The mechanics (if we may call it such) of the process is therefore that of daydreams, fantasy, pretending, perhaps, or whistling in the dark. A good deal of reasoning is of course involved, as in theology, but the conclusions at the end of it remain untested.

An obvious difficulty in the theory is that it cannot as it stands account for the fact that many of those who gain by the capitalist system are themselves religious devotees. If we are to believe the Marxists such people can have no motive for indulging in religious hopes, for their success in the real world provides them with no motive for compensating themselves in another. Are they, then, dispensers of opium instead of addicts? If so, we are back again with something very like what I earlier called the preliminary, superficial view of religion which Marxists ordinarily reject. The theory which

[9] 'God is the love which satisfies our wishes, the needs of our heart — he is the *realized wish* of the heart, the wish that has reached up to the certainty of its fulfilment, of its validity, to an undoubting certainty before which no contradiction on the part of the understanding, no objection from experience or the external world can remain standing.' *The Essence of Christianity*, Chapter XIII.

[10] 'The power of miracles', he says in Chapter XIV of *The Essence of Christianity*, 'is nothing but the power of the imagination.' 'By means of miracles', he goes on, 'ends are achieved without means.'

Marx and Engels based upon Feuerbach's writings must therefore be supplemented if it is to be consistent with these facts.

It may be said that religious beliefs and practices which originally sprang from the unsatisfied desires of all or most men continue to command the support of a prosperous minority as a result of tradition or irrational loyalty. This can be reinforced by the suggestion that the minority would have no strong motive for attacking a system of beliefs and practices which in fact supported their interests. There is a big difference, it can be argued, between consciously and cynically upholding a belief which one considers false and merely being impervious to the flaws in beliefs which one has been brought up to accept and which work to one's advantage. But it will be seen that now quite a different type of explanation is being provided for the religious beliefs of the bourgeoisie from that provided for those of proletarians. For whereas the latter (and men of the pre-scientific era too) are held to be engaged in what we have described as a sort of day-dreaming, the bourgeois is just prejudiced in favour of a tradition which fosters his interests. There is some similarity, of course, for both are prompted by their wishes to hold beliefs which do not stand up to testing. But whereas the proletarian dreams out imaginary substitutions for real satisfactions, the bourgeois joins in the dreaming to his own material advantage. Perhaps he defends religious belief because he dimly senses its value to himself, and thus he may employ the device known as rationalization — the manufacture of reasons for conduct which in any case one would engage in independently of the reasons so manufactured. The bourgeois who dimly feels that the continuance of religious belief is to the advantage of his class may thus become a religious apologist or idealist philosopher in order to rationalize his interests. Proletarian religion, therefore, is explained in terms of compensatory day-dreaming, bourgeois religion in terms of rationalization. These different psychological mechanisms operate within a system of social conflicts.

(*d*) I mentioned earlier that according to Feuerbach when religious people worship God they are, unawares, bowing down before something they themselves have created. Something for which they themselves are responsible has, so to say, the appearance of being alien or strange to them. Feuerbach's approach to religion was, as we should say today, psychological. He thought that dreams and insanity threw light on the nature of religious experience.[11] When Feuerbach speaks of God as man objectified, and when he talks of

[11] 'Religion is the dream of the human mind. But even in dreams we are not nowhere or in heaven, but on earth — in the kingdom of reality, only we are not looking at real things in the light of reality and necessity, but in the ravishing appearance of imagination and wilfulness.' Preface of the second edition of *The Essence of Christianity*.

the alienation and self-estrangement of man in his religious feelings, he seems to be thinking not merely in terms of daydreams but in terms of hallucinations. At any rate he is not thinking in terms of social institutions. Marx, I suggest, was thinking in terms of social institutions when he wrote his so-called Paris Manuscripts of 1844. These manuscripts, it should be emphasized, were a draft of a book to be called *Zur Kritik der Nationalökonomie: mit einem Schlusskapitel über die Hegelsche Philosophie* (Towards a Critique of Political Economy: with a Concluding Chapter on the Hegelian Philosophy). Marx brings together in these unfinished writings comments on classical economic theory and criticisms of Hegel along lines commenced by Feuerbach. Here he transfers the word 'estrangement' from the religious and philosophical sphere to industrial society where, he holds, the individual worker is in the grip of social arrangements which are inhuman and alien to him. The creations that have their creators in their grip are not dreams or hallucinations, but social institutions. Men lose their real way in pursuit of money just as in the past they thought that the fetish in the stone idol would bring them prosperity. Under the capitalist arrangements a man's own labour becomes estranged from him. 'The harder the worker works', writes Marx, 'the more powerful becomes the strange, objective world which he creates over against himself, and the poorer he himself becomes, and his inner world, and all the less there is that is his own. It is just the same in religion. The more man puts into God the less he has for himself . . . The alienation of the worker in his product does not only mean that his work becomes an object, an external existent, but that it exists *outside* him, independent of and strange to him, and becomes a self-existent power over against him, and the life he has transmitted to the object opposes him as something hostile and strange.'[12] Marx does not mean that the actual physical products of a man's own labour threaten him, but that the very purpose of work gets lost sight of and men become means instead of ends when large-scale industry produces goods under a capitalist organization. It is not money that governs men, but false beliefs about it and the social system within which it is used and pursued.

Now I am not here concerned with the truth of Marx's account of capitalism, but with the type of explanation that is involved in it. The idea is that men unwittingly produce social arrangements which inhibit the potentialities of most of them. This is only superficially like the idea that men create the gods and then serve their own creations, for the gods are not real existences which could be really

[12] *MEGA*, 1, 3, pp. 83-4.

served, whereas there are joint stock companies which employ workmen for wages. Marx seems to ignore religious institutions altogether. Indeed, he speaks of religion as 'the spiritual aroma'[13] of the imperfect society in which we live. It is clear that when he uses this phrase he is thinking in terms of imaginations and day-dreams which arise because of the failure of economic institutions to bring general satisfaction. What he does not consider are religious institutions themselves. Yet the religious beliefs which we in fact have are generally acquired through such institutions. Churches and priesthoods cannot be understood entirely in terms of compensations or even of rationalizations. We see here Marx's inadequacy as a sociologist – important social institutions are dismissed from consideration or regarded as ineffective because they are thought of as being like a ghostly smell. That is, Marx looks upon religious ideologies as nothing but a sort of shadow or distorted image of economic or industrial realities. He thus loses sight of religious institutions, or at any rate assumes that they are only secondary. But even if they are secondary, they are as real as those institutions which Marx considers primary. There is nothing shadowy, etherial or hallucinatory about them. They have whatever reality any other human institution has.

(*e*) What is it then to be in the grip of institutions which we ourselves have created? Marx himself gives an account of this when he discusses the division of labour in the *German Ideology*. With the division of labour, he says, 'man's own deed becomes an alien power opposed to him, which enslaves him instead of being controlled by him'.[14] Marx here appears to mean that in the very process of increasing their productive power by means of the division of labour, the individuals are forced into narrow grooves from which they cannot escape. These are not physical grooves. What prevents escape from them is the need to earn one's living. Other people are earning their living in other grooves and may not welcome intruders from outside. These are social, not physical grooves, yet when an individual gets into one of them it is as though he had jumped into a pit with no way up again. The longer he is there earning his bread by his specialized skill, the less likely it is that he will ever get out. Each individual gets his livelihood by working at an occupation which is prescribed to him in a system of interrelated occupations. No one has organized the whole system – this has developed from the interplay of competing individual producers. Thus in Vol. I of

[13] 'The struggle against religion is therefore indirectly the struggle against that world whose spiritual aroma religion is.' *MEGA*, I, I, i, p. 607.

[14] *MEGA*, I, 5, p. 22.

Capital Marx writes of human beings who 'behave atomistically in the extant social process of production' and says that in consequence 'the material form of their productive relations is independent of their own control and of their conscious individual activities. . .' He contrasts with these atomistic individuals of the market economy the co-operating members of a family in a domestic economy or the members of primitive communist groups. He even asks us to imagine what it would be like for the operations on Robinson Crusoe's island to be carried out by a Robinson Crusoe Syndicate which distributed the product in proportion to the hours worked by each of the members. His idea seems to be that there is something simple and intelligible about such arrangements, however defective they may be when measured by results. In such communities men at any rate know what is going on. When they have to pay feudal dues in forced labour they know how many hours they have worked, and cannot fail to realize that they are under the dominion of others. But they don't know what is going on in a developed market economy. The economy is now much more productive, but the individual workers cannot grasp its mysterious operations.

Now it seems to me that there are three different points that need to be distinguished.

(i) Any human society must surely have certain modes of behaviour prescribed for its members — modes of behaviour which last on from one generation to another, whether 'generation' is taken in a biological sense or more widely as with the successions of entrants to a school or club. No single individual can do very much to change these modes, and it is likely that he will be so trained that he will not wish to do so. The very process of training the young to enter the social order subjects them to social arrangements which neither they nor their teachers have created. Thus later generations are necessarily caught up in some institutions for which they are in no way responsible. To this extent, alienation is and must be a feature of any enduring society.

(ii) When new ideas about how to live arise in different parts of a society and are put into practice there, the results, because of the impact of these lives upon one another, may be quite different from what any of the originators intended. This is certainly what happened when large-scale production in factories for sale in a world market became prevalent. The workings of this system are not easy to understand, and the disadvantages for many were, when Marx was writing, at least as obvious as the advantages to some others. Social changes now occurred which seemed to be as unavoidable as the institutions of a static society. It was therefore reasonable to ask

whether enough could be discovered about how these things happened to enable the undesirable features of the unintended occurrences to be diminished or eradicated. I suggest that Marx was more confident than he should have been about what it was possible to do. For when people are free to invent and contrive, and when they do this from centres which are independent of one another, then the outcome of their unconcerted efforts will in many cases be unpredicted and, in practice if not in principle, be unpredictable. This does not mean that particular evils cannot be removed. But it does mean that the results of removing an evil, or a group of evils, may be quite unknown. It is thus impossible to combine all the benefits of tradition with all the benefits of progress. We may say, therefore, that not only alienation, but even some measure of estrangement, is a necessary feature of any *free* society.

(iii) It will have been noticed that Marx thought that a communist society would be unmysterious. We should all know what is going on in it and not be surprised or puzzled or made to wonder. The relations between men will be as obvious as feudal domination but at the same time rational and conscious. 'The religious reflection of the real world,' writes Marx in *Capital*, vol. I, p. 54, 'can only disappear when the relations between human beings in their practical everyday life have assumed the aspect of perfectly intelligible and reasonable relations as between man and man, and as between man and nature. The life-process of society, this meaning the material process of production, will not lose its veil of mystery until it becomes a process carried on by a free association of producers, under their conscious and purposive control.'

The whole section of *Capital* (Vol. I, Part I, Chapter 1, Section 4 – 'The Mystery of the Fetichistic Character of Commodities') from which these sentences are quoted gives a most significant indication of the social ideal by which Marx was attracted. In the ideal society there will no longer be any mystery, but the relations between men will be 'perfectly simple, perfectly clear'. There will be nothing ambiguous about them. In the *Critique of Political Economy* Marx had spoken with scorn of 'Robinsonades' as 'insipid illusions of the eighteenth century'. The reason for his scorn, in this context, was the idea of considering economic activity from a purely individual rather than from a fundamentally social standpoint. But in the section of *Capital* from which I have just quoted Marx uses Robinson Crusoe's methods of economy to illustrate a simple, clear and rational economic order. What appears to have attracted Marx, apart from the clarity and simplicity of it all, is that Robinson Crusoe had no hostile, possibly disturbing wills to take into account. We may,

indeed, remember his terror when at length he saw Man Friday's footprint in the sand. 'I stood like one thunderstruck, or as if I had seen an apparition . . . When I came to my castle (for so I think I called it for ever after this) I fled into it like one pursued . . . never frightened hare fled to cover or fox to earth with more terror than I to this retreat.' He was no longer, like Alexander Selkirk, the original of Robinson Crusoe, in Cowper's poem, monarch of all he surveyed.[15] He had someone else to reckon with, some unknown will to arouse perplexity and fear. In Marx's island there are no Man Fridays, since all its inhabitants are united into a single Robinsonian will which has nothing to contend with except the forces of nature. Nature, of course, can inspire men with fear, but not with the sort of fear that is associated with an unfathomed will. For the men of Marx's collectivist society will not believe that there is will or purpose in or behind natural processes. Collective mankind will be like a sort of divinity which has created a social world in which there is no devil to frustrate its purposes. There will be no conscious or unconscious conflicts in the individual or in the collectivity. The whole of mankind will form a common front against nature. There will be difficult problems, of course, but no really baffling or unanswerable ones. Like the Logical Positivists of our own time, Marx believed that all genuine problems can be solved. What is vague and shifting will be brought into focus in due time. Mystery, he held, is always a sign of confusion or dishonesty, never of profundity.

[15] William Cowper, *The Solitude of Alexander Selkirk.*

I am monarch of all I survey,
My right there is none to dispute;
From the centre all round to the sea,
I am lord of the fowl and the brute.

But Selkirk's conclusion was:

Better dwell in the midst of alarms
Than reign in this horrible place.

ON LEŚNIEWSKI'S ONTOLOGY[1]

By Czesław Lejewski

The study group which before the Second World War became known as the Warsaw School of Logic, was, throughout its only too short history, dominated by philosophers rather than by mathematicians.[2] Although Łukasiewicz and Leśniewski, the two founders of the School, held chairs in the Faculty of Mathematics and Natural Sciences in the University of Warsaw, neither of them had a science degree or a degree in mathematics. Both were Arts scholars or, to be more exact, both had graduated in philosophy under Twardowski in Lwów. They had, of course, studied mathematics but within the province of logic philosophy seems to have inspired their research. Łukasiewicz missed no opportunity to acknowledge his debt to Aristotle and Leśniewski meant his logic to be a true, though very general, description of reality, a kind of πρώτη φιλοσοφία. It is significant that the Warsaw School of Logic influenced the development of philosophy in Poland to an extent to which it never succeeded in influencing the development of mathematics.

Both Łukasiewicz and Leśniewski were masters in logical precision but they did not indulge in 'pure' formalism. This is particularly true of Leśniewski. He used to mock 'mathematical games' which had nothing to offer except the possibility of transforming certain initial formulae into other formulae in accordance with somewhat conventional rules. Contrary to the fashion prevailing among mathematicians Leśniewski insisted that only *true* propositions should be allowed as axioms of a deductive theory and that only those rules of transformation should be admitted which embodied intuitively *valid* rules of inference. He did not reject formalization as a method of presenting deductive theories. Far from it. But he maintained that a theory which was intuitively sound, did not lose its meaning through

[1] An earlier version of this paper was read and discussed at a logical colloquium held in Oxford in July 1956.

I am indebted to Mr. F. P. Chambers, Mr. W. C. Kneale, Professor Julius Kraft, Mr. E. C. Luschei, Professor A. N. Prior, Father Ivo Thomas, and Professor H. J. Woodger for their helpful suggestions which enabled me to improve the composition of my essay.

I should also like to express my gratitude to Dr. B. Sobociński for his kind permission to include some of his own unpublished results in the present paper, and for his generous comments, from which I have benefited a great deal.

[2] For a detailed discussion of the revival of logical and philosophical studies in Poland before the war see Bibliography, Ajdukiewicz [1], and Jordan [3]. See also Sobociński [16].

being subjected to the process of formalization. He consequently described himself as a confirmed intuitionist who at the same time was a radical formalist, and a better characterization of his standpoint could hardly be given.[3]

Leśniewski's criticism of 'pure' formalism shows that he had never ceased to be a philosopher. There are many 'pure' formalists among logicians and mathematicians but there are few 'pure' formalists among philosophers. For philosophers are, for the most part, preoccupied with the problem of meaning. Whether they deal with expressions of ordinary language or with logical formulae, they are concerned with interpretation rather than with formal elegance alone. The doctrine of 'pure' formalists could perhaps be condensed into the following motto: formalization before interpretation. Leśniewski's principle would read in the reverse.

For the most part Leśniewski's published papers present his theories at the stage of formalization with the problems of interpretation either left out or touched upon in an incidental manner. This makes the reading of these papers extremely difficult. It is the aim of the present contribution to bring the problems of interpretation to the foreground and by so doing serve as an informal introduction to one of the principal theories conceived by Leśniewski.

Leśniewski's System of Logic

Leśniewski's system of logic consists of two theories, which he called Protothetic and Ontology. Of these the former is perhaps the most comprehensive Logic of Propositions which has ever been devised. It goes beyond the classical Calculus of Propositions in several respects. It allows for instance for functorial variables for which constant functors of the Calculus of Propositions can be substituted.[4] It provides for the use of the universal quantifier to bind both the propositional and the functorial variables. It has a rule of definitions, which enables us to extend at will the variety of semantical categories within the field of the Logic of Propositions,

[3] See Leśniewski [8], p. 78.

[4] One of the simplest protothetical theses has the form of the following expression:

A. $$[f p]{:}f(p).\equiv .f(f(f(p)))$$

In this thesis 'p' is a propositional variable while 'f' is a functorial variable of the same semantical category as, for instance, the functor of negation '$\sim$'. Thus from A we get

B. $$[p]{:}\sim(p).\equiv .\sim(\sim(\sim(p)))$$

by applying the rule of substitution.

For the explanation of the symbolism see p. 160 f. In the terminology of the Warsaw School of Logic the term 'functor', which is due to Kotarbiński, denotes both propositional connectors, like 'if', 'or', 'and', and operators, like '+', '×' in mathematics. We characterize the semantical category of a functor by saying what it forms, whether a proposition or a name, how many arguments it requires, whether one, two or more, and to what semantical category these arguments belong.

and, in addition, it has a rule of extensionality; but the most significant point about Protothetic is that with its aid we can derive theses which enable us to dispose of the usual rules for operating with the universal quantifier in any deductive theory of lower generality.[5] In the edifice of the possible deductive theories Protothetic forms the very base. It requires no more fundamental theory than itself whereas other deductive theories, not included in it, have to be built on it or on a part of it. This is the case with Ontology.

If Protothetic is the most comprehensive Logic of Propositions then Ontology is the most comprehensive Logic of Names. Roughly speaking, it comprises the traditional logic in its modernized form and has counterparts of the Calculus of Predicates, the Calculus of Classes, and the Calculus of Relations including the Theory of Identity.[6]

The title of this paper indicates that Ontology will be our main concern. Protothetic will of course be presupposed but we shall make no explicit use of protothetical theses in any deductions that may be required in the course of this study. This means that the proofs of ontological theorems will be based on 'intuition' rather than on explicit references to various laws of the Logic of Propositions. In this respect our proofs will not differ from the proofs which we ordinarily find in mathematics.

At first sight Leśniewski's use of the term *Ontology* may seem curious and daring, but it should become clearer as we proceed that he was eminently justified. In fact his whole conception of logic was ontological through and through in a truly classical sense.

Although Ontology was the subject of several university courses given by Leśniewski during the twenty years of his academic career in Warsaw until his death in 1939, there have been few papers published on it. The fundamental and most authoritative source is Leśniewski's 'Über die Grundlagen der Ontologie', *Comptes rendus des séances de la Societé des Sciences et des Lettres de Varsovie*, Classe III, XXIII Année, Warszawa 1930, pp. 111-32. It is an extremely condensed and difficult paper as it was meant to be a sort of 'identity card' of Ontology and not its 'lengthy biography'. It gives an axiom of Ontology and, with reference to pp. 59-67 of Leśniewski's 'Grundzüge eines neuen Systems der Grundlagen der Mathematik', *Fundamenta Mathematicae*, vol. 14, Warszawa 1929, pp. 1-81, it also gives the rules of inference for Ontology stated here with a precision which has not since been improved upon. In addition, the paper

[5] In this connection see Leśniewski [8], p. 38 f.

[6] According to Leśniewski himself, Ontology is nearest to Schröder's Calculus of Classes provided the latter includes the theory of individuals. See Leśniewski [8], p. 5.

contains a brief account of the researches of Leśniewski and his collaborators into the axiomatic foundations of Ontology.[7]

There are two more papers by Leśniewski on special problems theoretically belonging to Ontology. They are 'Über Funktionen, deren Felder Gruppen mit Rücksicht auf diese Funktionen sind', *Fundamenta Mathematicae*, vol. 13, Warszawa 1929, pp. 319-33, and 'Über Funktionen, deren Felder Abelsche Gruppen in Bezug auf diese Funktionen sind', *Fundamenta Mathematicae*, vol. 14, Warszawa 1929, pp. 242-51. These papers are also worth mentioning for the fact that they contain some of the neatest examples of Leśniewski's method of setting out his deductions.

Finally, there is 'O podstawach matematyki, Rozdział XI' (On the Foundations of Mathematics, Chapter XI), *Przegląd Filozoficzny*, vol. 34, Warszawa 1931, pp. 153-70, where Leśniewski gives his analysis of the meaning of the primitive constant of Ontology as used in the 'Grundlagen der Ontologie'. Naturally enough this analysis is made from the point of view of the Polish language.

This is all that Leśniewski himself ever published on his Ontology. His copious notes and manuscripts, which contained a wealth of new results and which were to have been prepared for publication by Sobociński, were destroyed in 1944 during the war.

There have been some references to Ontology by other authors. In the Bibliography appended to this paper the relevant contributions are indicated with an asterisk.

Contrary to the widely accepted practice Leśniewski intended his logic to be an interpreted system. He attached definite meanings to his constants and regarded the theses of his deductive theories as true propositions in the sense in which propositions of empirical sciences are accepted as true. It is with this in mind that we should approach his theories. Ontology has been described as the most comprehensive Logic of Names because its most characteristic expressions belong to the semantical category of names, just as the most characteristic expressions of Protothetic belong to the semantical category of propositions. If, however, we take into account the contents of Ontology then it would be more appropriate to describe it as a theory of what there is. Just as astronomy tells us something about heavenly bodies, the theses of Ontology tell us something about things, or objects if one prefers, or individuals. Since in accordance with Leśniewski's intentions theses of Ontology are to be regarded not as mere 'well formed formulae' but as meaningful propositions which can be examined for their truth or falsity, it is essential that we should understand as clearly as possible the ontological vocabulary.

[7] A detailed report on these researches is contained in Sobociński [14].

Only when we have mastered this vocabulary can a further step be made, namely the one which consists of a critical study of those ontological theses which are already at our disposal. Finally, we may try to discover new truths, which we can formulate in terms of the ontological vocabulary, and assign them their proper places in the theory by deducing them from the axiom or establishing their independence.

The Ontological Table

Learning the ontological vocabulary, like learning any other vocabulary, is an informal affair. Any method is good as long as it gives the required results. The meaning of ontological constants might be best explained by translating them into English. The ambiguity of ordinary usage, however, renders this method less effective than we might like it to be, and compels us to have recourse to other devices. In the logic of propositions the method of determining the meaning of the various proposition-forming functors for propositional arguments with the aid of truth tables has proved to be very helpful. In what follows we shall try to construct an 'Ontological Table' and with its aid indicate the meaning of certain ontological constants in a way that should leave less room for misunderstanding than a direct translation.

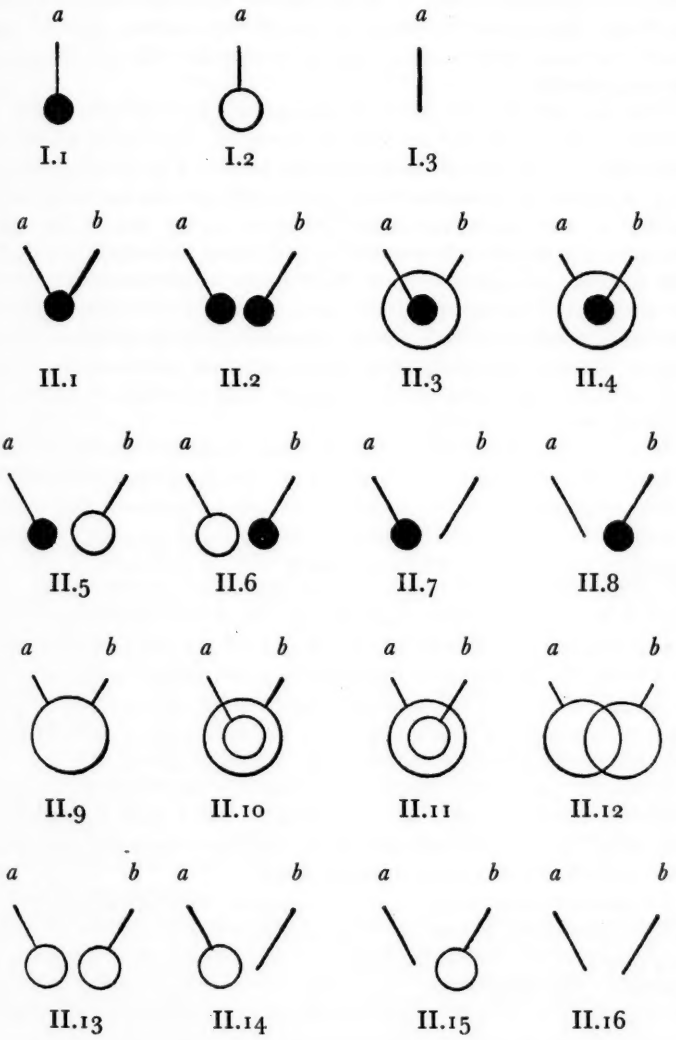
The Ontological Table we have in mind is an extension of the well-known Eulerian diagrams. Among the names and name-like expressions of ordinary language we can distinguish (1) unshared names, each of which names one object only, *e.g.* 'Socrates', 'the Moon', *etc.*, (2) shared names, each of which names more than one object, *e.g.* 'satellite of Jupiter', 'heavenly body', *etc.*, and (3) fictitious names, *i.e.*, expressions which as regards their syntax behave like unshared or shared names, but which do not name anything, *e.g.* 'Pegasus', 'centaur', 'object which does not exist.'[8]

The Eulerian diagrams, the purpose of which was to illustrate the meaning of the syllogistic expressions, do not take into account unshared names or fictitious names. Thus the idea calls for extension. In the Ontological Table below we shall use a shaded circle to represent the only object named by an unshared name, and an unshaded one to represent the many objects each of which is named by a shared name. No circle will be used in the case of a fictitious name. Thus the diagrams I.1, I.2, and I.3 below will serve to illustrate what one might call the semantical status of single names.

[8] The terminology 'unshared names', 'shared names' and 'fictitious names' is due to J. H. Woodger. See his *Biology and Language*, Cambridge, 1952, p. 17 and 'Science without Properties', *The British Journal for the Philosophy of Science*, II (1952), p. 196.

The diagrams numbered from II.1 to II.16 will illustrate the semantical status of pairs of names. Analogous diagrams could be easily worked out for sets consisting of three or more names.

Ontological Table



It should not be difficult to grasp what is meant when we say that diagram I.1 in the Ontological Table illustrates the semantical status of the name 'the Moon', or that diagram I.3 illustrates the semantical status of the name 'centaur'. The shaded circle in I.1 reminds us of the fact that there is exactly one object named by the name 'the Moon'. I.3 has no circle at all because the names whose semantical status it is meant to illustrate, name nothing existing. The name 'centaur' is one of such names. Of the name 'man' we can correctly say that its semantical status is illustrated by diagram I.2.

The case of pairs of names is equally simple. Diagram II.3, for instance, illustrates the semantical status of such pairs as the one consisting of 'the Moon' and 'heavenly body'. The shaded circle in II.3 is inside the unshaded one to remind us that the only object named by the name 'the moon' happens to be one of the many objects each of which is named by the name 'heavenly body'. To take another example, diagram II.12 illustrates the semantical status of such pairs of names as the one consisting of 'painter' and 'sculptor'. In this case the circles — which are unshaded to indicate that the names 'painter' and 'sculptor' name several individuals each — intersect to show that there are individuals each of which is named by both the names involved.

Now the usefulness of the Ontological Table lies in this: with its aid we can determine the meaning of certain proposition-forming functors for one or for two arguments which are names. And what is more, we can do so with a fairly high degree of precision. Strictly speaking, the functors whose meaning will be determined with the aid of the Ontological Table do not occur in ordinary usage, just as for instance the word 'infant' does not occur in ordinary language in the sense in which it is used by the Local Education Authorities. In a sense the functors of Ontology are 'technical' terms like the technical terms used by lawyers, stockbrokers, or engineers. They serve the purpose of expressing one's thoughts with more precision than can be attained by means of the ordinary idiom.

Primarily ontological functors are used in their graphical form. However, for the purpose of reading them out I shall equate them with their nearest counterparts from ordinary language, and hope that this will cause no misunderstanding.

The functors selected for characterizing the ontological vocabulary will be introduced under the following headings: functors of inclusion, functors of exclusion, functors of identity, and functors of existence. We begin with

(a) *Functors of inclusion*

1. *The functor of singular inclusion* occurs in expressions of the type '$a \in b$' (to be read: *a* is *b*). A proposition of this type will be said to be true if and only if diagram II.1 or diagram II.3 illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

2. *The functor of strong inclusion* occurs in expressions of the type '$a \sqsubset b$' (to be read: every *a* is *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.1, II.3, II.9, or II.10, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

3. *The functor of weak inclusion* occurs in expressions of the type '$a \subset b$' (to be read: all *a* is *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.1, II.3, II.8, II.9, II.10, II.15, or II.16, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

4. *The functor of partial inclusion* occurs in expressions of the type '$a \triangle b$' (to be read: some *a* is *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.1, II.3, II.4, II.9, II.10, II.11, or II.12, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

(b) *Functors of exclusion*

5. *The functor of singular exclusion* occurs in expressions of the type '$a \notin b$' (to be read: *a* is-not *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.2, II.5, or II.7, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

6. *The functor of strong exclusion* occurs in expressions of the type '$a \not\sqsubset b$' (to be read: every *a* is-not *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.2, II.5, II.6, II.7, II.13, or II.14, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

7. *The functor of weak exclusion* occurs in expressions of the type '$a \not\subset b$' (to be read: no *a* is *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.2, II.5, II.6, II.7, II.8, II.13, II.14, II.15, or II.16, illustrates the

semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

8. *The functor of partial exclusion* occurs in expressions of the type '$a \not\triangle b$' (to be read: some *a* is-not *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.2, II.4, II.5, II.6, II.7, II.11, II.12, II.13, or II.14, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

(c) *Functors of identity*

9. *The functor of singular identity* occurs in expressions of the type '$a = b$' (to be read: *a* is the same object as *b*). A proposition of this type will be said to be true if and only if diagram II.1 illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

10. *The functor of strong identity* occurs in expressions of the type '$a \square b$' (to be read: only every *a* is *b*). A proposition of this type will be said to be true if and only if diagram II.1 or diagram II.9 illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.[9]

11. *The functor of weak identity* occurs in expressions of the type '$a \bigcirc b$' (to be read: *a* is identical with *b*). A proposition of this type will be said to be true if and only if one of the following diagrams, II.1, II.9, or II.16, illustrates the semantical status of the two names for which the '*a*' and the '*b*' stand. Otherwise it will be a false proposition.

(d) *Functors of existence*

12. *The first functor of existence* occurs in expressions of the type '$\mathrm{ex}(a)$' (to be read: there exists at least one *a*). A proposition of this type will be said to be true if and only if diagram I.1 or diagram I.2 illustrates the semantical status of the name for which the '*a*' stands. Otherwise it will be a false proposition.

13. *The second functor of existence* occurs in expressions of the type '$\mathrm{sol}(a)$' (to be read: there exists at most one *a*). A proposition of this type will be said to be true if and only if diagram I.1 or diagram I.3 illustrates the semantical status of the name for which the '*a*' stands. Otherwise it will be a false proposition.

14. *The third functor of existence* occurs in expressions of the type '$\mathrm{ob}(a)$' (to be read: there exists exactly one *a*). A proposition of this type will be said to be true if and only if diagram I.1 illustrates the

[9] I have included this functor in the list on the suggestion of Dr. Sobociński.

semantical status of the name for which the '*a*' stands. Otherwise it will be a false proposition.

Here we end our list of ontological functors. It is quite clear that with the aid of the Ontological Table many more functors could be introduced. The ones I have included in the list appear to be of especial interest. Thus the functor of singular inclusion plays the part of the only undefined constant in Leśniewski's original system of Ontology. Propositions formulated with the aid of the functor of strong inclusion and those formulated with the aid of the functor of partial inclusion have the same meaning as the meaning of the two affirmative categorical propositions of syllogistic, the universal and the particular respectively, provided we agree to regard the singular affirmative proposition as a special case of either. The functor of weak inclusion has the meaning which we sometimes attribute to the functor of inclusion in Boolean Algebra.

Among the functors of exclusion we do not find functors which could be said to form universal negative or particular negative propositions of syllogistic. A negative proposition of syllogistic has 'existential import' with respect to both its arguments, *i.e.*, with respect to the subject and with respect to the predicate. This means that in a true negative proposition of syllogistic neither the subject nor the predicate can be a fictitious name. Now a true proposition formed with the aid of any of the functors of exclusion above distinguished may have a fictitious name as the second argument, *i.e.*, as the predicate, and the functor of weak exclusion produces a true proposition if the first argument, *i.e.*, the subject, is a fictitious name. It may be of interest to note that a weak exclusion is the contradictory of the corresponding partial inclusion and that a partial exclusion is the contradictory of the corresponding weak inclusion.

The distinction between the various functors of identity appears to be very useful, particularly in connection with the problem whether identity implies existence or not. It seems to be evident from what has been said that if by identity we mean weak identity then there is no ground for asserting that it implies existence. On the other hand the inference from identity to existence is quite legitimate if by identity we mean strong identity or singular identity.

The functors of existence will prove very useful in formulating simple ontological theses which eventually will lead to the formulation of axioms.

Theses relating different Functors

Now that we know the meaning of a number of ontological

functors, we can proceed to constructing a few propositions about what there is. In doing so we shall have at our disposal the vocabulary of the Logic of Propositions including the vocabulary of the Theory of Quantification. For as it has already been said, Ontology presupposes Protothetic.

We shall be using the Peano-Russellian symbolism for the purpose of formulating ontological theses and it may not be out of place to give a very brief account of it with a view to meeting our present requirements only. Small letters of the initial portion of the Latin alphabet will serve as variables for which any name, unshared, shared, or fictitious, can be substituted. In contradistinction to *propositional* variables, like *p*, *q*, *r*, . . ., they may conveniently be described as *nominal* variables. Together with appropriate functors they form propositional functions of which the following may serve as examples: $a \in b$, $a \sqsubset c$, $d \not\subset b$, $b = a$, $c \notin d$, ex(*a*), sol(*a*), *etc.* Out of these simple propositional functions compound propositional functions can be formed by using proposition-forming functors for propositional arguments. Thus for instance we have '$a = b. \equiv .b = a$' to mean '*a* is the same object as *b* if and only if *b* is the same object as *a*', '$a \in b. \supset .b \triangle a$' to mean 'if *a* is *b* then some *b* is *a*', '$a \in b. \vee .a \notin b$' to mean '*a* is *b* or *a* is-not *b*', '$a \in b.b \in a$' to mean '*a* is *b* and *b* is *a*', and finally '$\sim(a \in b)$' to mean 'it is not true that *a* is *b*'. In addition to proposition-forming functors for propositional arguments we shall make use of the quantifiers: the universal quantifier '[*a b c* . . .]' to mean 'for all *a*, *b*, *c*, . . .' and the particular quantifier '[$\exists$ *a b c* . . .]' to mean 'for some *a*, *b*, *c*, . . .'. Accordingly, the expressions '$[a].a \bigcirc a$' and '$[\exists a].\sim(a \in a)$' mean 'for all *a*, $-a$ is identical with *a*' and 'for some *a*, — it is not true that *a* is *a*'.

The use of dots as a 'punctuation' device is one of the characteristic features of the symbolism and calls for explanation.[10] If a dot or a cluster of dots adjoins, *i.e.*, immediately precedes or immediately follows one of the following '$\equiv$', '$\supset$', or '$\vee$', then it only indicates the scope of the functor. Thus if in an expression we find a cluster of say three dots immediately preceding '$\supset$' then this indicates that the antecedent of the implication involved extends to the left until it reaches a cluster of more than three dots, or failing that until it reaches the beginning of the expression under consideration. Analogously, a cluster of three dots immediately following '$\supset$' indicates that the consequent of the implication involved extends to the right up to the first cluster of more than three dots, or failing that up to the end of the expression under

[10] See A. N. Whitehead and B. Russell, *Principia Mathematica*, vol. I, Cambridge 1935, p. 94.

consideration. If in an expression a dot or a cluster of say five dots immediately follows a quantifier then it indicates that the expression 'covered' by the quantifier extends to the right over the clusters of less than five dots, and over the clusters of five dots provided the latter are not immediately followed by '$\equiv$', '$\supset$', or '$\vee$', until it reaches a cluster of five dots immediately followed by '$\equiv$' or '$\supset$' or '$\vee$', or — failing that — until it reaches a cluster of more than five dots. If neither is the case, the quantifier covers the whole of the expression to its right. A dot or a cluster of say four dots which neither adjoins '$\equiv$' or '$\supset$' or '$\vee$' nor immediately follows a quantifier means 'and', *i.e.*, it stands for the functor of conjunction. The first argument of the conjunction extends to the left of the cluster over the clusters of less than four dots until it reaches a cluster of four or more dots. The second argument of the conjunction is indicated in an analogous way, and the procedure is generalized to apply to conjunctions with more than two conjuncts.

In order to leave as few doubts as possible as regards the interpretation of the symbolism, a number of ontological theses will be given below together with a translation into ordinary language.

If we bear in mind the meaning we have attached to the various ontological functors then little reflection will be needed to convince ourselves of the truth of the following theses:

T1. $[a]:\mathrm{ex}(a). \equiv .[\exists b].b \in a$
i.e., for all *a*, — there exists at least one *a* if and only if for some *b*, — *b* is *a*.

T2. $[a]:\mathrm{ex}(a). \equiv .[\exists b].b \sqsubset a$
i.e., for all *a*, — there exists at least one *a* if and only if for some *b*, — every *b* is *a*.

T3. $[a]:\mathrm{ex}\ (a). \equiv .[\exists b]. \sim (a \subset b)$
i.e., for all *a*, — there exists at least one *a* if and only if for some *b*, — it is not true that all *a* is *b*.

T4. $[a]:\mathrm{ex}(a). \equiv .[\exists b].a \not\sqsubset b$
i.e., for all *a*, — there exists at least one *a* if and only if for some *b*, every *a* is-not *b*.

T5. $[a]:.\mathrm{sol}(a). \equiv :[b\ c]:b \in a.c \in a. \supset .b \in c$
i.e., for all *a*, — there exists at most one *a* if and only if for all *b* and *c*, — if *b* is *a* and *c* is *a* then *b* is *c*.

T6. $[a]:.\mathrm{sol}(a). \equiv :[b]:b \in a. \supset .a \in b$
i.e., for all *a*, — there exists at most one *a* if and only if for all *b*, — if *b* is *a* then *a* is *b*.[11]

T7. $[a]:.\mathrm{sol}(a). \equiv :[b\ c]:b \sqsubset a.c \sqsubset a. \supset .b \sqsubset c$
i.e., for all *a*, — there exists at most one *a* if and only if for all *b* and *c*, — if every *b* is *a* and every *c* is *a* then every *b* is *c*.

T8. $[a]:.\mathrm{sol}(a). \equiv :[b]:b \sqsubset a. \supset .a \sqsubset b$
i.e., for all a, — there exists at most one a if and only if for all b, — if every b is a then every a is b.[11]

T9. $[a]::\mathrm{sol}(a). \equiv :.[b\ c\ d]:.b \subset a.c \subset a. \supset :b \subset c. \vee .c \subset d$
i.e., for all a, — there exists at most one a if and only if for all b, c, and d, — if all b is a and all c is a then (all b is c or all c is d).

T10. $[a]::\mathrm{sol}(a). \equiv :.[b\ c]:.b \subset a. \supset :a \subset b. \vee .b \subset c$
i.e., for all a, — there exists at most one a if and only if for all b and c, — if all b is a then (all a is b or all b is c).[11]

T11. $[a]:.\mathrm{sol}(a). \equiv :[b\ c]:b \triangle a.c \triangle a. \supset .b \triangle c$
i.e., for all a, — there exists at most one a if and only if for all b and c, — if some b is a and some c is a then some b is c.

T12. $[a]::\mathrm{sol}(a). \equiv :.[b\ c\ d]:.b \notin c.c \notin d. \supset :b \notin a. \vee .c \notin a$
i.e., for all a, — there exists at most one a, if and only if for all b, c, and d, — if b is-not c and c is-not d then (b is-not a or c is-not a).

T13. $[a]::\mathrm{sol}(a). \equiv :.[b\ c\ d]:.b \not\sqsubset c.c \not\sqsubset d. \supset :b \not\sqsubset a. \vee .c \not\sqsubset a$
i.e., for all a, — there exists at most one a if and only if for all b, c, and d, — if every b is-not c and every c is-not d then (every b is-not a or every c is-not a).

T14. $[a]::\mathrm{sol}(a). \equiv :.[b\ c]:.b \not\subset c. \supset :b \not\subset a. \vee .c \not\subset a$
i.e., for all a, — there exists at most one a if and only if for all b and c, — if no b is c then (no b is a or no c is a).

T15. $[a]:.\mathrm{sol}(a). \equiv :[b\ c]:b \not\triangle c.a \not\triangle b. \supset .b \not\triangle a$
i.e., for all a, — there exists at most one a if and only if for all b and c, — if some b is-not c and some a is-not b then some b is-not a.

T16. $[a]:\mathrm{ob}(a). \equiv .[\exists b].a \in b$
i.e., for all a, — there exists exactly one a if and only if for some b, — a is b.

T17. $[a]:\mathrm{ob}(a). \equiv .\mathrm{ex}(a).\mathrm{sol}(a)$
i.e., for all a, — there exists exactly one a if and only if (there exists at least one a and there exists at most one a).

The functor of singular inclusion suffices for the purpose of defining all the remaining functors whose meaning has been determined with the aid of the Ontological Table. This is evident from the following theses, which will now be given in symbols only.

T18. $[a\ b]::a \sqsubset b. \equiv :.[\exists c].c \in a:.[c]:c \in a. \supset .c \in b$
T19. $[a\ b]:.a \subset b. \equiv :[c]:c \in a. \supset .c \in b$
T20. $[a\ b]:a \triangle b. \equiv .[\exists c].c \in a.c \in b$

[11] T6, T8, and T10 are due to Sobociński.

T21. $[a\ b]{:}a \notin b. \equiv .a \in a. \sim (a \in b)$
T22. $[a\ b]{::}a \not\subset b. \equiv :.[\exists c].c \in a{:}.[c]{:}c \in a. \supset .\sim(c \in b)$
T23. $[a\ b]{:}.a \subset\!\!\!\!/\ b. \equiv :[c]{:}c \in a. \subset .\sim(c \in b)$
T24. $[a\ b]{:}a \not\subset b. \equiv .[\exists c].c \in a.\sim(c \in b)$
T25. $[a\ b]{:}a = b. \equiv .a \in b.b \in a$
T26. $[a\ b]{::}a \square b. \equiv :.[\exists c].c \in a{:}.[c]{:}c \in a. \equiv .c \in b$
T27. $[a\ b]{:}.a \bigcirc b. \equiv :[c]{:}c \in a. \equiv .c \in b$

Finally it may be worth noting that the following theses hold:

T28. $[a\ b]{:}a \sqsubset b. \equiv .\mathrm{ex}(a).a \subset b$
T29. $[a\ b]{:}a \not\subset b. \equiv .\mathrm{ex}(a).a \subset\!\!\!\!/\ b$
T30. $[a\ b]{:}a \subset\!\!\!\!/\ b. \equiv .\sim(a \triangle b)$
T31. $[a\ b]{:}a \not\subset b. \equiv .\sim(a \subset b)$
T32. $[a\ b]{:}a \square b. \equiv .\mathrm{ex}(a).a \bigcirc b$
T33. $[a\ b]{:}a = b. \equiv .\mathrm{sol}(a).a \square b$

Historical Review of Axioms of Ontology

Leśniewski's original system of Ontology constructed in 1920 was based on the following single axiom:

T34. $[a\ b]{::}a \in b. \equiv :.[\exists c].c \in a{:}.[c]{:}c \in a. \supset .c \in b{:}.\ [c\ d]{:}c \in a.d \in a. \supset .c \in d$

The meaning of T34 becomes clear as soon as we realize that in virtue of T1, T19, and T5 this thesis is inferentially equivalent to

T35. $[a\ b]{:}a \in b. \equiv .\mathrm{ex}(a).a \subset b.\mathrm{sol}(a)$

In 1921 Leśniewski found that T34 could be replaced as a single axiom of Ontology by

T36. $[a\ b]{::}a \in b. \equiv :.[\exists c].c \in a.c \in b{:}.[c\ d]{:}c \in a.d \in a. \supset .c \in d$

In virtue of T20 and T5, T36, which is shorter than T34, is inferentially equivalent to

T37. $[a\ b]{:}a \in b. \equiv .a \triangle b.\mathrm{sol}(a)$

On the basis of Protothetic alone we can easily show that T34 is inferentially equivalent to the following four theses:

T38. $[a\ b]{:}a \in b. \supset .[\exists c].c \in a$
T39. $[a\ b\ c]{:}a \in b.c \in a. \supset .c \in b$
T40. $[a\ b\ c\ d]{:}a \in b.c \in a.d \in a. \supset .c \in d$
T41. $[a\ b\ c]{::}c \in a{:}.[d]{:}d \in a. \supset .d \in b{:}.[d\ e]{:}d \in a.e \in a. \supset .d \in e{:}. \supset .a \in b$

In 1921 Tarski showed that by making use of an appropriate definition we can deduce T40 from T39, and in 1929 Sobociński succeeded in proving that T38 and T39 alone could serve as an axiom system of Ontology and that T41 could be deduced from them

by applying, in the course of deductions, the rule of extensionality, which at that time had already been added to Ontology. In the same year Sobociński established that

T42. $[a\ b]::a \in b. \equiv :.[\exists c].c \in a.c \in b:.[c]:c \in a. \supset .a \in c$

which is shorter than T36, could also be used as a single axiom of Ontology. In virtue of T20 and T6, T42 is inferentially equivalent to T37.

In connection with these results of Sobociński a further simplification of the axiomatic foundations of the system of Ontology with the functor of singular inclusion as a single undefined functor was achieved by Leśniewski, who in 1929 noticed that the axiom system consisting of T38 and T39 was inferentially equivalent to

T43. $[a\ b]:a \in b. \equiv .[\exists c].a \in c.c \in b$

which thus qualified to the status of single axiom of the theory.[12]

In 1946 I set myself the task of finding single axioms for ontological functors other than the functor of single inclusion. I remembered from Leśniewski's lectures that systems of Ontology could be constructed with the aid of the functor of strong inclusion or with the aid of the functor of partial inclusion, but no theses which could be used as the respective single axioms had ever come to my knowledge.

As a result of my research I found that the following thesis

T44. $[a\ b]::a \in b. \equiv :.a \sqsubset b:.[c\ d]:c \sqsubset a.d \sqsubset a. \supset .c \sqsubset d$

which in virtue of T7 is inferentially equivalent to

T45. $[a\ b]:a \in b. \equiv .a \sqsubset b.\mathrm{sol}(a)$

could be used as a definition of singular inclusion in terms of strong inclusion, and that a system of Ontology with the functor of strong inclusion as a single undefined functor could be based on the following axiom:

T46. $[a\ b]:::a \sqsubset b. \equiv ::.[\exists c].c \sqsubset a::.[c]::c \sqsubset a. \supset :.[\exists d]:.d \sqsubset c.d \sqsubset b:.$
$[e\ f]:e \sqsubset d.f \sqsubset d. \supset .e \sqsubset f$

It is not difficult to see from T2, T7, and T45 that T46 is inferentially equivalent to

T47. $[a\ b]::a \sqsubset b. \equiv :.\mathrm{ex}(a):.[c]:c \sqsubset a. \supset .[\exists d].d \in c.d \in b$

which thus throws some light on the meaning of T46.

As a further result I established that if we used

T48. $[a\ b]::a \in b. \equiv :.a \triangle b:.[c\ d]:c \triangle a.d \triangle a. \supset .c \triangle d$

as a definition of singular inclusion in terms of partial inclusion then

T49. $[a\ b]::a \triangle b. \equiv :.[\exists c]:.c \triangle a.c \triangle b:.[d\ e]:d \triangle c.e \triangle c. \supset$
$.d \triangle e$

[12] See Leśniewski [9] and Sobociński [14].

could be used as a single axiom for a system of Ontology with the functor of partial inclusion as a single undefined functor.

In virtue of T11, T48 is inferentially equivalent to

T50. $[a\ b]{:}a \in b. \equiv .a \triangle b.\mathrm{sol}(a)$

and T49 is inferentially equivalent to T20 in virtue of T11 and T50.

In addition I found that the functor of singular exclusion could be used as a single undefined functor in a system of Ontology. In such a system

T51. $[a\ b]{:}a \in b. \equiv .[\exists c].a \notin c.\sim(a \notin b)$

could be used as a definition of singular inclusion while the single axiom could have the form of the following expression:

T52. $[a\ b]{::}.a \notin b. \equiv ::[\exists c].c \notin b.\sim(c \notin a){::}[c\ d\ e]{:}.c \notin d.d \notin e. \supset :c \notin a. \vee .d \notin a$

The meaning of T52 can be seen from

T53. $[a\ b]{:}.a \notin b. \equiv :[\exists c].c \notin b.c \in a{:}\mathrm{sol}(a)$

which is inferentially equivalent to T52 in virtue of T51 and T12.

I returned to these and similar problems in 1953 when I found that a system of Ontology could be based on the functor of weak inclusion. Singular inclusion could be defined with the aid of the following thesis:

T54. $[a\ b]{::}.a \in b. \equiv ::[\exists c].\sim(a \subset c){:}a \subset b{::}[c\ d\ e]{:}.c \subset a.d \subset a. \supset :c \subset d. \vee .d \subset e$

which in virtue of T3 and T9 is inferentially equivalent to T35.

The following thesis:

T55. $[a\ b]{:::}a \subset b. \equiv ::.[c\ d]{::}.\sim(c \subset d).c \subset a. \supset ::[\exists\ e\ f]{::}\sim(e \subset f).e \subset c.e \subset b{::}[g\ h\ i]{:}.g \subset e.h \subset e. \supset :g \subset h. \vee .h \subset i$

could serve as a single axiom. We can easily grasp the meaning of this axiom if we compare it with

T56. $[a\ b]{:}.a \subset b. \equiv :[c]{:}c \sqsubset a. \supset .[\exists d].d \in c.d \in b$

which is inferentially equivalent to it in virtue of T3, T28, and T54.

In connection with this result I exchanged a number of letters with Dr. Sobociński, from whom I learned that Leśniewski had had axioms for the functors of strong inclusion, weak inclusion, and partial inclusion. Sobociński could not recollect what these axioms were. He remembered, however, that T54 was the definition used by Leśniewski. Owing to his own research Sobociński knew that T49 could be used as a single axiom for the functor of partial inclusion. It is more likely than not that this was also known to Leśniewski.

While analysing T55 Sobociński noticed that a system of Ontology with the functor of weak inclusion as a single undefined term could

be based on a shorter axiom, which had the form of the following expression:

T57. $[a\ b]:::a \subset b. \equiv ::. [cd]::.\sim(c \subset d).c \subset a::[d\ e\ f]:.d \subset c.e \subset c. \supset :d \subset e. \vee .e \subset f:: \supset .c \subset b$

In virtue of T54 this axiom is inferentially equivalent to

T58. $[a\ b]:.a \subset b. \equiv :[c]:c \in a. \supset .c \subset b$

A further simplification of the axiom for weak inclusion took place when Sobociński proved T10. This enabled him to shorten T54 as a definition of singular inclusion in terms of weak inclusion, and to replace T57 by the following thesis:

T59. $[a\ b]:::a \subset b. \equiv ::.[c\ d]::.\sim(c \subset d).c \subset a::[d\ e]:. d \subset c. \supset :c \subset d. \vee .d \subset e:: \supset .c \subset b$

which again is inferentially equivalent to T58 in virtue of T9, T10 and T54.

As a sequel to these results Sobociński proved T8 and

T60. $[a\ b]::a \in b. \equiv :.a \sqsubset b:.[c]:c \sqsubset a. \supset .a \sqsubset c$

which defines singular inclusion in terms of strong inclusion. He also established that the following thesis

T61. $[a\ b]:::a \sqsubset b. \equiv ::.[\exists c].c \sqsubset a::.[c]::c \sqsubset a:.[d]:d \sqsubset c. \supset .c \sqsubset d:. \supset .c \sqsubset b$

sufficed as a single axiom for the system of Ontology based on strong inclusion.

It is evident from T2 and T60 that T61 is inferentially equivalent to

T62. $[a\ b]::a \sqsubset b. \equiv :.\mathrm{ex}(a):.[c]:c \in a. \supset .c \sqsubset b$

At this stage Sobociński remarked that curiously enough T36 ,T49 and T61 — which are axioms for singular inclusion, partial inclusion, and strong inclusion respectively — consisted of six ontological units each.[13]

In 1956 I obtained further results concerning the axiomatic foundations of Ontology. I found that T59 could be replaced by the following shorter thesis:

T63. $[a\ b]:::a \subset b. \equiv ::.[c]::.c \subset a::[d\ e]:.d \subset c. \supset :c \subset d. \vee .d \subset e:: \supset .c \subset b$

This axiom consists of six ontological units and is inferentially equivalent to

T64. $[a\ b]:.a \subset b. \equiv :[c]:c \subset a.\mathrm{sol}(c). \supset .c \subset b$

as can be seen from T10.

[13] An ontological unit is a simple propositional function formed with the aid of an ontological functor.

I also found that systems of Ontology could be constructed on the basis of each of the remaining functors of exclusion. The theses given below can serve as the respective definitions of singular inclusion and as axioms:

T65. $[a\ b]::.a \in b. \equiv ::[\exists c].a \not\subset c:\sim(a \not\subset b)::[c\ d]:.c \not\subset d. \supset :a \not\subset c. \vee .a \not\subset d$

T66. $[a\ b]:::.a \not\subset b. \equiv :::[\exists\ c\ d].c \not\subset d.\sim(c \not\subset a):::[c\ d]::.\ c \not\subset d::[e\ f]:.e \not\subset f. \supset :c \not\subset e. \vee .c \not\subset f:: \supset :c \not\subset a. \vee .c \not\subset b$

T67. $[a\ b]::.a \in b. \equiv ::\sim(a \subsetneq b)::[c\ d]:.c \subsetneq d. \supset :c \subsetneq a. \vee .d \subsetneq a$

T68. $[a\ b]::.a \subsetneq b. \equiv ::[c]::[d\ e]:.d \subsetneq e. \supset :d \subsetneq c. \vee .e \subsetneq c:. \supset :c \subsetneq a. \vee .c \subsetneq b$

T69. $[a\ b]::a \in b. \equiv :.[\exists c].a \barwedge c.\sim(a \barwedge b):.[c\ d]:c \barwedge d.\ a \barwedge c. \supset .c \barwedge a$

T70. $[a\ b]::.a \barwedge b. \equiv ::[\exists c]::\sim(c \barwedge a).c \barwedge b:.[d\ e]:d \barwedge e.\ c \barwedge d. \supset .d \barwedge c$

It is interesting to note that T68, *i.e.*, the axiom for weak exclusion, and T70, *i.e.*, the axiom for partial exclusion, consist of six ontological units each. I have not, however, succeeded in finding six-unit axioms for singular exclusion or for strong exclusion.

The functor of singular exclusion deserves our attention because with its aid we can formulate a fairly simple frame for writing so-called ontological definitions. On adapting the rule of ontological definitions and the rule of extensionality to certain properties of singular exclusion I was able to prove that the following thesis:

T71. $[a\ c]:a \not\in b. \equiv .[\exists c\ d].a \not\in d.c \not\in b.\sim(c \not\in a)$

could be used as a single axiom of a system of Ontology. This result dates from 1954.[14]

Proof that a system of Ontology can be based on partial Inclusion

My summary of the results obtained in connection with the study of the axiomatic foundations of Ontology is far from being exhaustive. In particular it gives no insight into the character and amount of deductive considerations which have to be completed before a thesis can be declared to rank as a single axiom of a system of Ontology based on a single undefined functor other than the functor of singular inclusion. In what follows I propose to show how it can be established that T49, for instance, can be used as a single axiom of the system of Ontology with the functor of partial inclusion as the only undefined term. The problem is solved by proving that such a

[14] For the purpose of establishing the adequacy of the axioms for '$\subsetneq$', '$\not\in$' '$\not\subset$', and '$\barwedge$' use is made of ontological definitions, whose nature is discussed in the last section of this paper.

system is inferentially equivalent to the original system of Ontology based on singular inclusion. Since the functor of singular inclusion and the functor of partial inclusion are two different functors we have to deduce T49 and T48, which is a definition of singular inclusion in terms of partial inclusion, from T36 and T20, *i.e.*, from a possible axiom of the original system of Ontology and from a definition of partial inclusion in terms of singular inclusion. Then we have to show that T36 and T20 follow from T49 and T48. To put it in self-explanatory symbols we have to establish that

$$\{\text{T36, T20}\} \rightleftarrows \{\text{T49, T48}\}$$

Our deductions proceed as follows:

Axiom:

A. $=$ T36. $[a\ b]::a \in b. \equiv :.[\exists\ c].c \in a.c \in b:.[c\ d]:c \in a.d \in a. \supset .c \in d$

Definition:

D. $=$ T20. $[a\ b]:a\ \triangle\ b. \equiv .[\exists\ c].c \in a.c \in b$

TT1. $[a\ b\ c]:a \in b.c \in a. \supset .c \in b$

Proof: $[a\ b\ c]::$

(1) $a \in b.$

(2) $c \in a. \supset :.$

(3) $[d\ e]:d \in a.e \in a. \supset .d \in e:.$ (follows from A and 1)

(4) $a \in a.$ (from A, 2, 3)

(5) $a \in c:.$ (3, 4, 2)

(6) $[d\ e]:d \in c.e \in c. \supset .d \in e:.$ (A, 2)

$c \in b$ (A, 5, 1, 6)

TT2. $[a\ b]:a\ \triangle\ b. \supset .b\ \triangle\ a$ (D)

TT3. $[a\ b]:a \in b. \supset .a\ \triangle\ b$

$[a\ b]:.$

(1) $a \in b. \supset :$

(2) $[\exists\ c].c \in a.c \in b:$ (A, 1)

$a\ \triangle\ b$ (D, 2)

TT4. $[a\ b\ c\ d]:a \in b.c\ \triangle\ a.d\ \triangle\ a. \supset .c\ \triangle\ d$

$[a\ b\ c\ d]::$

(1) $a \in b.$

(2) $c\ \triangle\ a.$

(3) $d\ \triangle\ a. \supset :.$

$[\exists\ e]:$

(4) $e \in c.$
(5) $e \in a:$
$[\exists f].$
(6) $f \in d.$ } (D, 2)
(7) $f \in a.$ } (D, 3)
(8) $e \in f:$ (A, 1, 5, 7)
(9) $e \in d:.$ (TT1, 6, 8)
$c \triangle d$ (D, 4, 9)

TT5. $[a\ b\ c]:a \in b.b \in c. \supset .b \in a$
$[a\ b\ c]::$
(1) $a \in b.$
(2) $b \in c. \supset :.$
(3) $[d\ e]:d \in b.e \in b. \supset .d \in e:.$ (A, 2)
(4) $b \in b:.$ (A, 1, 3)
$b \in a$ (3, 4, 1)

TT6. $[a\ b\ c]::[d\ e]:d \triangle a.e \triangle a. \supset .d \triangle e:.b \in a.c \in a:. \supset .b \in c$
$[a\ b\ c]::$
(1) $[d\ e]:d \triangle a.e \triangle a. \supset .d \triangle e:.$
(2) $b \in a.$
(3) $c \in a:. \supset :$
(4) $b \triangle a.$ (TT3, 2)
(5) $c \triangle a.$ (TT3, 3)
(6) $b \triangle c:$ (1, 4, 5)
$[\exists d].$
(7) $d \in b.$ } (D, 6)
(8) $d \in c.$ } (D, 6)
(9) $b \in d:$ (TT5, 7, 2)
$b \in c$ (TT1, 8, 9)

TT7. $[a\ b]::a \triangle b:.[c\ d]:c \triangle a.d \triangle a. \supset .c \triangle d:. \supset .a \in b$
$[a\ b]::$
(1) $a \triangle b:.$
(2) $[c\ d]:c \triangle a.d \triangle a. \supset .c \triangle d:. \supset :.$
(3) $[c\ d]:c \in a.d \in a. \supset .c \in d:.$ (TT6, 2)
(4) $[\exists c].c \in a.c \in b:.$ (D, 1)
$a \in b$ (A, 4, 3)

TT8. $= T48.$ $[a\ b]::a \in b. \equiv :.a \triangle b:.[c\ d]:c \triangle a.d \triangle a. \supset .$
$c \triangle d$[15] (TT3, TT4, TT7)

TT9. $[a\ b]::a \triangle b. \supset :.[\exists c]:.c \triangle a.c \triangle b:.[d\ e]:d \triangle c.e \triangle c. \supset$
$.d \triangle e$
$[a\ b]::.$

[15] In Słupecki [13] this thesis is deduced from T34

(1) $a \triangle b. \supset ::$
$[\exists c]:.$
(2) $c \in a.$
(3) $c \in b.$ } (D, 1)
(4) $c \triangle a.$ (TT3, 2)
(5) $c \triangle b:.$ (TT3, 3)
(6) $[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e::$ (TT4, 2)
$[\exists c]:.c \triangle a.c \triangle b:.[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e$ (4, 5, 6)

TT10. $[a\ b\ c]::c \triangle a.c \triangle b:.[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e:. \supset .a \triangle b$
$[a\ b\ c]::$
(1) $c \triangle a.$
(2) $c \triangle b:.$
(3) $[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e:. \supset .$
(4) $a \triangle c.$ (TT2, 1)
(5) $b \triangle c.$ (TT2, 2)
$a \triangle b$ (3, 4, 5)

TT11. $= T49.$ $[a\ b]::a \triangle b. \equiv :.[\exists c]:.c \triangle a.c \triangle b:.[d\ e]:d \triangle c.$
$e \triangle c. \supset .d \triangle e$ (TT9, TT10)

It is evident from TT11 and TT8 that T49 and T48 are deducible as theses within the framework of the original system of Ontology. It remains to show that T36 and T20 can be derived from T49 and T48, if the latter theses are regarded as the axiom and a definition respectively.

Axiom:

TTA. $=$ TT11 $[a\ b]::a \triangle b. \equiv :.[\exists c]:.c \triangle a.c \triangle b:.[d\ e]:d \triangle c.$
$e \triangle c. \supset .d \triangle e$

Definition:

TTD. $=$ TT8 $[a\ b]::a \in b. \equiv :.a \triangle b:.[c\ d]:c \triangle a.d \triangle a. \supset .c \triangle d$

TT11.1 $[a\ b]:a \triangle b. \supset .[\exists c].c \in a.c \in b$
$[a\ b]::.$
(1) $a \triangle b. \supset ::$
$[\exists c]:.$
(2) $c \triangle a.$
(3) $c \triangle b:.$
(4) $[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e:.$ } (TTA, 1)
(5) $c \in a.$ (TTD, 2, 4)
(6) $c \in b::$ (TTD, 3, 4)
$[\exists c].c \in a.c \in b$ (5, 6)

TT11.2 $[a\ b\ c]:c \in a.c \in b. \supset .a \triangle b$
$[a\ b\ c]::$
(1) $c \in a.$
(2) $c \in b. \supset :.$
(3) $c \triangle a.$ (TTD, 1)
(4) $c \triangle b:.$
(5) $[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e:.$ } (TTD, 2)
$a \triangle b$ (TTA, 3, 4, 5)

TT11.3 = T20 $[a\ b]:a \triangle b. \equiv .[\exists\ c].c \in a.c \in b$ (TT11.1, TT11.2)

TT11.4 $[a\ b]:a \in b. \supset .[\exists\ c].c \in a.c \in b$
$[a\ b]:.$
(1) $a \in b. \supset :$
(2) $a \triangle b:$ (TTD, 1)
$[\exists\ c].c \in a.c \in b$ (TT11.1, 2)

TT11.5 $[a\ b\ c\ d]:a \in b.c \in a.d \in a. \supset .c \in d$
$[a\ b\ c\ d]::$
(1) $a \in b.$
(2) $c \in a.$
(3) $d \in a. \supset :.$
(4) $[e\ f]:e \triangle a.f \triangle a. \supset .e \triangle f:.$ (TTD, 1)
(5) $c \triangle a:.$ } (TTD, 2)
(6) $[e\ f]:e \triangle c.f \triangle c. \supset .e \triangle f:.$
(7) $d \triangle a.$ (TTD, 3)
(8) $c \triangle d:.$ (4, 5, 7)
$c \in d$ (TTD, 8, 6)

TT11.6 $[a\ b\ c]::[d\ e]:d \in a.e \in a. \supset .d \in e:.b \triangle a.c \triangle a:. \supset .b \triangle c$
$[a\ b\ c]:::$
(1) $[d\ e]:d \in a.e \in a. \supset .d \in e:.$
(2) $b \triangle a.$
(3) $c \triangle a:. \supset ::.$
$[\exists\ d]::$
(4) $d \triangle b.$ } (TTA, 2)
(5) $d \triangle a:.$
(6) $[e\ f]:e \triangle d.f \triangle d. \supset .e \triangle f:.$
(7) $d \in a::$ (TTD, 5, 6)
$[\exists\ e]:.$
(8) $e \triangle c.$ } (TTA, 3)
(9) $e \triangle a:.$
(10) $[f\ g]:f \triangle e.g \triangle e. \supset .f \triangle g:.$
(11) $e \in a.$ (TTD, 9, 10)
(12) $e \in d.$ (1, 11, 7)

(13) $e \triangle d::$ (TTD, 12)
(14) $d \triangle c::.$ (TTA, 13, 8, 10)
$b \triangle c$ (TTA, 4, 14, 6)

TT11.7 $[a\ b\ c]::c \in a.c \in b:.[d\ e]:d \in a.e \in a. \supset .d \in e:. \supset .a \in b$

$[a\ b\ c]::$
(1) $c \in a.$
(2) $c \in b:.$
(3) $[d\ e]:d \in a.e \in a. \supset .d \in e:. \supset :.$
(4) $[d\ e]:d \triangle a.e \triangle a. \supset .d \triangle e:.$ (TT11.6, 3)
(5) $c \triangle a.$ (TTD, 1)
(6) $c \triangle b:.$ } (TTD, 2)
(7) $[d\ e]:d \triangle c.e \triangle c. \supset .d \triangle e:.$ }
(8) $a \triangle b:.$ (TTA, 5, 6, 7)
$a \in b$ (TTD, 8, 4)

TT11.8 = T36 $[a\ b]::a \in b. \equiv :.[\exists\ c].c \in a.c \in b:.[d\ e]:d \in a.\ e \in a. \supset .d \in e$ (TT11.4, TT11.5, TT11.7)

From TT11.8 and TT11.3 we see that T49 and T48 imply T36 and T20. This completes the proof that a system of Ontology inferentially equivalent to the original system can be based on T49 as a single axiom with the functor of partial inclusion as a single undefined form.

Definitions in Systems of Ontology

Now that we have a suitable axiom for the functor of partial inclusion it remains to adapt the rule for writing ontological definitions to the properties of this functor.

There are two types of definition in Ontology: protothetical or propositional definitions on the one hand and ontological or nominal definitions on the other. Propositional definitions are called so because the most characteristic functors introduced by them are proposition-forming functors. The most characteristic functors introduced by means of nominal definitions are name-forming functors. Without any claim to precision the corresponding two rules can be briefly characterized as follows:[16]

(*a*) *The rule of protothetical definitions.* On the assumption that at the given stage in the development of a system of Ontology a thesis T is the last thesis, the rule of protothetical definitions entitles us to add to the system as a new thesis an expression of the type:

[16] In the present formulation many-link functions are not taken into consideration because no definitions involving such functions are required for the purpose of establishing the adequacy of the axioms discussed in this paper, except in the case of T42, T43, T71 and T77. For the authoritative statement of the rules of definition for the system based on T43 see Leśniewski [9].

(i) $$[\ldots] \,.\, \alpha \equiv \beta$$

provided the following conditions are fulfilled: (1) 'α' stands for a propositional function. The functor of this function is a constant term which does not occur in T or in any of the theses preceding T. The arguments of the function are all variables each of which is of the same semantical category as an expression occurring in T or in a thesis preceding T. Variables which are of the same semantical category as certain expressions obtainable exclusively within the logic of propositions are not allowed in 'α'. All the variables in 'α' are free in 'α' and no variable in 'α' occurs more than once; (2) 'β' stands for a propositional expression. Every constant in 'β' belongs to the logic of propositions or occurs in T or in at least one thesis preceding T. Every variable in 'β' is of the same semantical category as an expression occurring in T or in a thesis preceding T. Every variable occurring in 'α' occurs in 'β' as a free variable in 'β'. Every free variable in 'β' occurs in 'α'. No variable is free in (i).

It is easy to see that in accordance with this rule any of the expressions: T1, T5, T6, T16, T − 18T27, could be added as a definition to the system of Ontology with T34, T36, T42 or T43 as the axiom. It would suffice to assume that the axiom was the last thesis of the system. But we would have to assume that T1 and T5 were already in the system if we wanted to add T17 as a definition. The system of Ontology based on T49 as the axiom could be immediately extended by adding T11, T30, and T48 as definitions.

(*b*) *The rule of ontological definitions.* On the assumption that at the given stage in the development of the original system of Ontology a thesis T is the last thesis, the rule of ontological definitions enables us to add to the system as a new thesis an expression of the type:

(ii) $$[a \ldots]{:}a \in \Psi. \equiv .a \in a.\varphi(a)$$

provided the following conditions are fulfilled: (1) 'Ψ' stands for a constant name, which does not occur in T or in a thesis preceding T, or it stands for a nominal function. If the latter is the case then the functor of this function is a constant term which does not occur in T or in any of the theses preceding T, while the arguments of the function are all variables each of which is of the same semantical category as an expression occurring in T or in a thesis preceding T. Variables which are of the same semantical category as certain expressions obtainable exclusively within the logic of propositions are not allowed in 'Ψ'. All the variables in '$a \in \Psi$' are free in '$a \in \Psi$' and none of them occurs in '$a \in \Psi$' more than once. (2) '$\varphi(a)$' stands for a propositional expression. Every constant in '$\varphi(a)$' belongs to the logic of propositions or occurs in T or in at least one thesis

preceding T. Every variable in '$\varphi(a)$' is of the same semantical category as an expression occurring in T or in a thesis preceding T. Every variable occurring in '$a \in \Psi$' occurs in '$a \in a.\varphi(a)$' as a free variable. Every free variable occurring in '$a \in a.\varphi(a)$' occurs in '$a \in \Psi$'. No variable is free in (ii).

The following theses are examples of expressions which can be added as ontological definitions to the system based on T36 as the axiom:

T72. $[a]{:}a \in \mathsf{V}. \equiv .a \in a$
i.e., for all a, — a is an object if and only if a is a,

T73. $[a]{:}a \in \Lambda. \equiv .a \in a.\sim(a \in a)$
i.e., for all a, — a is an object which does not exist if and only if (a is a and it is not true that a is a),

T74. $[a\ b]{:}a \in \mathsf{N}(b). \equiv .a \in a.\sim(a \in b)$
i.e., for all a and b, — a is non-b if and only if (a is a and it is not true that a is b).

The rule of protothetical definitions remains the same for any system of Ontology irrespective of which functor is used as the undefined ontological term. This, however, does not hold for the rule of ontological definitions. Schema (ii) above represents what might be called the definitional frame for a system with the functor of singular inclusion as the primitive term. A different definitional frame is required for a system based on a functor other than the functor of singular inclusion.

A trivial way of obtaining such a frame would be to 'translate' schema (ii) with the aid of the definition of '$\in$' in terms of the functor involved. But the frame secured in this manner proves to be long and clumsy. Thus the problem arises of finding a shortest definitional frame for a given functor.

Since the following thesis:

T75. $[b\ \varphi]{::}.[a]{:}a \in b. \equiv .a \in a.\varphi(a){:}. \equiv {::}[a]{::}a \bigtriangleup b. \equiv {:}.[\exists\ c]{:}.c \bigtriangleup a{:}.[d\ e]{:}d \bigtriangleup c.e \bigtriangleup c. \supset .d \bigtriangleup e{:}.\varphi(c)$

can be proved in Ontology, we conclude that, in the case of the system based on T49 as the axiom, the definitional frame for ontological definitions can take the form of the following schema:

(iii) $[a \ldots]{::}a \bigtriangleup \Psi. \equiv {:}.[\exists\ c]{:}.c \bigtriangleup a{:}.[d\ e]{:}d \bigtriangleup c.e \bigtriangleup c. \supset .d \bigtriangleup e{:}.\varphi(c)$

On the assumption that at the given stage in the development of the system a thesis T is the last thesis, an expression of type (iii) can be added to the system as a definition provided 'Ψ' and '$\varphi(c)$' satisfy the conditions on the lines of those discussed under (b) above. In

particular it is required that every variable occurring in 'Ψ' should occur in '$\varphi(c)$' as a free variable, and that every free variable in '$\varphi(c)$' should occur in 'Ψ' or be equiform with 'c'.

Thus in the system under consideration the functor '$\mathcal{N}$' could be defined as follows:

T76. $[a\ b]::a\ \Delta\ \mathcal{N}(b).\ \equiv\ :.[\exists\ c]:.c\ \Delta\ a:.[d\ e]:d\ \Delta\ c.e\ \Delta\ c.\supset .d\ \Delta\ e:.\sim(c\ \Delta\ b)$

It is not unlikely that by making full use of the rule of ontological definitions in the system based on partial inclusion, and by exploiting the rule of extensionality, whose adaptation to the system could not be discussed in this paper, we may obtain a simpler axiom or a simpler frame for ontological definitions.

In connection with the rule of ontological definitions for systems with the functor of singular inclusion as the primitive term I was anxious to make clear that in the formulation given the rule presupposed T34 or T36 as the axiom. Interestingly enough this formulation proves to be too weak if we want to use T43 as the axiom.[17] T43 has many attractions as a single axiom. This is why Leśniewski eventually decided in favour of strengthening the rule of ontological definitions, which he formulated with the aid of the following frame:

(iv) $[a\ \ldots].a\ \varepsilon\ \Psi.\ \equiv .\varphi(a)$

In this schema '$\varphi(a)$' is of the form '$a\ \varepsilon\ \Omega$' or it is a conjunction which contains at least one conjunct of the form '$a\ \varepsilon\ \Omega$'. It is not difficult to see that an expression which is a definition in accordance with schema (ii) is also a definition in accordance with schema (iv). But the converse does not hold.

If we wanted to retain the weaker rule of ontological definitions then T43 would have to be replaced by

T77. $[a\ b]:.a\ \varepsilon\ b.\ \equiv\ :a\ \varepsilon\ a:[\exists\ c].a\ \varepsilon\ c.c\ \varepsilon\ b$

This was established by Sobociński in 1936.

BIBLIOGRAPHY

[1] K. Ajdukiewicz, 'Der logische Antiirrationalismus in Polen', *Erkenntnis*, vol. V, 1935.

*[2] K. Ajdukiewicz, 'Die syntaktische Konnexität', *Studia Philosophica*, vol. I, Lwów 1935.

*[3] Z. Jordan, *The Development of Mathematical Logic and of Logical Positivism in Poland between the Two Wars*, London 1945.

[17] Dr. Sobociński drew my attention to this fact.

*[4] T. KOTARBIŃSKI, *Elementy teorji poznania, logiki formalnej i metodologji nauk* (Elements of Epistemology, Formal Logic and Methodology), Lwów 1929.

*[5] C. LEJEWSKI, 'Logic and Existence', *The British Journal for the Philosophy of Science*, vol. V, Edinburgh 1954.

[6] S. LEŚNIEWSKI, 'Über Funktionen, deren Felder Gruppen mit Rücksicht auf diese Funktionen sind', *Fundamenta Mathematicae*, vol. XIII, Warszawa 1929.

[7] S. LEŚNIEWSKI, 'Über Funktionen, deren Felder Abelsche Gruppen in Bezug auf diese Funktionen sind', *Fundamenta Mathematicae*, vol. XIV, Warszawa 1929.

[8] S. LEŚNIEWSKI, 'Grundzüge eines neuen Systems der Grundlagen der Mathematik', *Fundamenta Mathematicae*, vol. XIV, Warszawa 1929.

[9] S. LEŚNIEWSKI, 'Über die Grundlagen der Ontologie', *Comptes rendus des séances de la Societé des Sciences et des Lettres de Varsovie*, Classe III, XXIII Année, Warszawa 1930.

[10] S. LEŚNIEWSKI, 'O podstawach matematyki', Rozdział XI' (On the Foundations of Mathematics, Chapter XI), *Przegląd Filozoficzny*, vol. XXXIV, Warszawa 1931.

*[11] J. ŁUKASIEWICZ, 'The Principle of Individuation', Aristotelian Society Proceedings, *Supplementary Volume XXVII*, London 1953.

*[12] A. N. PRIOR, *Formal Logic*, Oxford 1955.

*[13] J. SŁUPECKI, 'S. Leśniewski's Calculus of Names', *Studia Logica*, vol. III, Warszawa 1955.

*[14] B. SOBOCIŃSKI, 'O kolejnych uproszczeniach aksjomatyki "ontologji" prof. St. Leśniewskiego' (On Successive Simplifications of the Axiom-system of Leśniewski's 'Ontology'), *Księga Pamiątkowa — Fragmenty Filozoficzne*, Warszawa 1934.

*[15] B. SOBOCIŃSKI, 'L'analyse de l'antinomie Russellienne par Leśniewski', *Methodos*, vol. I and vol. II, Milano 1949 and 1950.

[16] B. SOBOCIŃSKI, 'In Memoriam Jan Łukasiewicz', *Philosophical Studies*, vol. VI, Maynooth (Ireland) 1956.

REVIEWS

System of Ethics by LEONARD NELSON. Translated by Norbert Guterman. Foreword by H. J. Paton. Introduction by Julius Kraft. New Haven: Yale University Press (London, Oxford University Press), 1956. Pp. xxi + 285. $5.

One can hardly imagine a book which a journal called *Ratio* could more appropriately review. Except Spinoza's *Ethics*, I know of no book in ethical literature in which the appeal is so consistently and exclusively to reason. Nelson proceeds from a set of basic principles – the validation of which he carried out in his *Critique of Practical Reason* (1917) – and deduces from them an extensive theory of duties and ideals, introducing empirical considerations only when he has to, and even then preserving the deductive form of exposition. The philosopher to whom he says he owes most is Jacob Fries, but Fries was a by-product of Kant, and any attentive reader will be constantly reminded of the rationalism of the *Metaphysics of Morals* and the second *Critique*.

Who was Leonard Nelson? To present-day philosophers he is an almost unknown name. He is little read in Germany, virtually unheard of in Britain, and only beginning to be known in the United States. One of his loyal former students, L. H. Grunebaum, has instituted in New York a Leonard Nelson Foundation, devoted to bringing his work to attention, and has offered generous prizes to American students for the best essays on his thought. The Yale University Press has recently published two of his books, *Socratic Method and Critical Philosophy* and the work now under review. To each of these volumes an introduction has been written by another of Nelson's devoted students, Julius Kraft, and the earlier of these introductory essays, containing a helpful sketch of Nelson's life and work, has been reprinted here as an appendix. A foreword is contributed by the distinguished Kantian scholar H. J. Paton, formerly professor of moral philosophy at Glasgow and Oxford, who owns that he has been attracted, impressed, and indeed excited, by such writing of Nelson as he has read, and regrets not having made his acquaintance earlier.

The man who has aroused this interest was, despite his English name, a German, born in 1882, the son of a prominent Berlin lawyer, and related through his mother to the composer Mendelssohn. He was educated at the Universities of Berlin, Heidelberg and Göttingen, and for the last eighteen years of his life taught philosophy at Göttingen as lecturer and professor. He developed a Socratic method of teaching that seems to have left a deep impression on his students. But his interests ranged far beyond philosophy. He was an intimate friend of the mathematician Hilbert, to whom one of his books is dedicated, and is known to symbolic logicians as co-author of the Nelson-Grelling paradox about autological words. He

was also engrossed in politics and courageous in political activity. Opposed to nationalism, whether German or any other, he publicly defended international peace and the limitation of sovereignty in the midst of the First World War, and wrote a book, *Rechtswissenschaft ohne Recht*, to show that the positivist theory of law, a legal relativism which would put existing law beyond appeal, is jurisprudence without justice. He founded the Walkemühle School near Cassel, largely to train students in enlightened and liberal citizenship. Falling under the ban of Hitler, the school was moved to Denmark and later to England. There is no doubt that Nelson himself would have fallen under that ban if he had lived. But he died in 1927 at the age of 45.

He left behind him a mass of writing, much of it still in manuscript. His history of metaphysics has still to see the light. Of a large-scale work on *The Foundations of Ethics* two volumes appeared before his death, the first entitled *Critique of Practical Reason* in 1917, and the third entitled *System of the Philosophy of Law and Politics* in 1924. The middle volume, *System of Philosophical Ethics and Pedagogics*, was published posthumously in 1932. The present book is a translation, excellently done, of the first part of this volume, including the ethics but omitting the section on educational theory.

Nelson's mode of attack on every philosophical issue was determined by his theory of knowledge. He was the most uncompromising of rationalists. He held that if we go back far enough in our thinking in any field, we shall find a set of principles which are presupposed and are therefore fundamental, but are themselves incapable of proof. Natural science, for example, assumes that every event is connected by law with some preceding event. This principle cannot be proved deductively, and any inductive argument for it presupposes its truth. What is the status of such principles? Some philosophers have supposed them to be analytically true; Mill thought they were very wide empirical generalizations; most philosophers perhaps regard them as postulates which we cannot demonstrate, but which all experience confirms. Nelson's view differed from all these; to him they were synthetic *a priori* insights. They were not, like the corresponding Kantian principles, merely ways of ordering sense experience; the reason which achieved them was laying bare the structure of the real world. And the first task of philosophy is analysis, not merely in the positivist sense of defining terms, but in the Socratic sense of bringing to light these major premises and assumptions on which all reasoning in a particular field is based. The second task of philosophy is to develop systematically the implications of these principles. 'A philosophical science deserving of the name,' says Nelson, 'is always a whole in which no part is arbitrary . . . For in a pure philosophical science the structure and the articulation of the whole are clearly prescribed by reason.'

Now ethics is in this sense a philosophical science. To be sure, it does not start, as metaphysics does, from propositions about what is, for it deals with what ought to be, and Nelson is clear that unless 'ought' appears in the premises, the conclusion will never be able to extract it. But he is

clear also that ethics is not, as Kant supposed, merely a theory of duties. 'We should do not only our duty but more than that, namely all that is worth doing, which deserves to be done, even if we are not duty-bound to do it.' I may amuse myself this evening with a comic book or I may read Nelson; neither may present itself as a duty, but I may still profit more by doing the latter. Ethics has a larger scope than morality. It should consider not only duties, but also those ideals whose fulfilment means culture. The joint achievement of these ends Nelson describes as wisdom.

Still, duty is the prime concern of ethics, and he devotes the first part of his book to developing the theory of duties. The initial step is to get clear what duty means. Unfortunately it is indefinable, since it defies analysis into anything simpler. But at least we can say of it what Kant did, that it is a categorical, not a hypothetical imperative. To do an act as a means to some good for oneself or another is not to do one's duty; indeed it is not to act morally at all. To act morally is to do one's duty just because it is one's duty; and to see one's duty in a particular case is to grasp a law requiring anyone so situated to act in a certain way. This is involved analytically, Nelson would say, in the very idea of duty.

He now moves on to what is involved in this idea synthetically, though still necessarily. We can see that the moral law is not merely a concept; it is an imperative that binds in practice. It is not real in the sense in which the law of gravitation is real, for that would mean that there are in fact no exceptions to it. It still holds, though constantly violated. This can only mean that there is more in moral conduct than can be explained by natural law; when acting from the sense of duty, I may override what psychology would account my strongest impulse; the rational insight that leads one to act is not determined by psychological law, nor is the process of leading which it initiates. The will, so far as rational, is free. And 'character' means the dominance of the sense of duty over the impulses of the natural man. Nelson is a modern Stoic. 'The stronger the passions, the more pronounced the spontaneity of character; for it is measured precisely by the maximum strength of the passions which the will can master.' The man of character is the reasonable man, licensing and repressing his feelings, forming habits and breaking them, at the behest of principle, rationally disclosed. Since duty is acting from the grasp of principle, and such grasp is the function of reason, 'moral truthfulness, i.e. the striving for moral clarity' is an imperative that is universal and formally deducible.

But how are we to know what reason requires of us in a particular case? Mere principle is not enough. We must pass over, therefore, into a material theory of duties and take into account the circumstances of the case. But what circumstances? Nelson replies, the *interests* of the persons affected. From this follows a number of further principles, even in advance of our knowing what in particular these circumstances may be. There is the principle of personal dignity, which requires respect for the interests of others. There is the principle of equality: 'each person *per se* has equal dignity with every other person'. There is the principle of equity, which

requires that we should treat the interests of others as of equal importance with like interests of our own.

Two of Nelson's deductions in 'the material theory of duties' are especially arresting. He holds that it is a grave mistake to regulate punishment by utilitarian considerations. The deliberate violation of a moral law calls for retribution to the guilty agent, and 'this retribution consists in his suffering an injury to his interests equal to the one he has unrightfully inflicted on others'. If there are readers who find in such teaching a touch of cruelty, they may well consider a second deduction. Nelson holds that although only rational beings can have duties, any being that has interests has rights that are based on those interests. Now animals clearly have interests, the interest in living, for example, and in not suffering pain. It follows that animals have rights against men; Nelson devotes a special chapter to establishing this. Through the interstices of his iron deduction one can feel at times the glow of a warm heart.

Ethics can thus lay down many principles that restrict the conduct of the moral man. But it can give no advice for the particular case. Only one who knows the interests affected can do that.

Having deduced the main principles of duty, Nelson turns to the theory of ideals, or of goods, and attempts a similar process there. For him the pursuit of goods and of moral goodness, of culture and character, are two quite different things, and yet the principles of the two pursuits can be dealt with in parallel ways. Hence Nelson starts again from the abstract concept of an ideal or good, and works out its implications. If anything is really good, it must be so for all men of clear vision; it will be better for each of us that such clear vision be developed; we shall be cultivated persons only to the extent to which we see and are governed by these ideals and so on. Indeed the implications of an ideal are similar to those of duty except for one large difference. The claim of duty is always paramount. Hence there is no room in morality for tolerance; a man can always do his duty, and if he does not, there is no good excuse for him. But there are all degrees of approximation to a goal, and full achievement of it may be beyond us. Hence in the pursuit of ideals tolerance must be the rule.

What are these ideals to be? If we consult our desires, we shall find plenty of answers, but Nelson insists that one cannot infer what we ought to want from what we do in fact want. There must be some means of discriminating legitimate goals from illegitimate ones. What this is will surprise no one. It is rationality. Ideals must prove satisfactory to persons; a person is a set of impulses which a rational will is seeking to bring under control; the ends of these impulses must therefore justify themselves at the bar of reason. Three ideals pass the test. They are the love of truth, the love of justice, and the love of beauty. Nelson believes that each of these can be derived and justified by a process of strict though synthetic reasoning, and that each can be shown to require the others.

Nelson's discussion of how these ideals are misunderstood and misapplied in practice, though it does not, of course, have the rigour of deduction, is to my mind the most interesting part of his book. It is rich in the

sort of insights that could be achieved only by a keen intelligence turned in earnest reflection upon the human scene. The quality of them may be seen in a few examples. 'Resignation ought to be man's predominant mood . . . A man who makes himself hostage to the vagaries of fortune voluntarily surrenders his freedom and remains at best a slave to fortune who merely happens to be lucky' (178-9). 'Willingness to compromise is a test of a man's idealism' (177). 'The ideal of originality is . . . the poorest and most wretched of all, for it is defined by mere otherness' (194). 'The development of theoretical knowledge is in itself a technical vocation, and if it does not permeate life, if it does not contribute to the rule of reason over life, it produces only skilful scholars' (205). 'Culture is rationality of life . . .' (213). 'No one can be an object of pure love unless he loves the truth' (217). 'Love and hate are correlated . . . Unless there is hate of the ugly, there can be no love of the beautiful . . .' (219). 'Enlightenment is sometimes condemned on the ground that it surrenders life to the cold and prosaic rule of the understanding. But the warmth we are supposed to win by renouncing reflection is a false warmth. Doubt and analysis mark man's first steps towards autonomous activity, and hence towards culture . . . respect for truth is the very foundation of culture' (235). These passages suggest the quality of Nelson's mind – austere, intellectually conscientious and exacting, unsentimental, independent in thought, concise in utterance.

His book presents so complete a contrast to current ethical discussions that one can hardly believe it appeared as recently as the 'thirties. That was a time when the 'emotivism' of Russell and the Vienna Circle was just beginning to make itself felt. If Nelson was writing today, would he have put his case differently? In one respect I hope he would. Empiricism has restored the fashion of discussing principles through particular cases. A sound rationalism need have no fear of such cases, and if it is to avoid the charge of confining itself, like a sputnik, to the stratosphere, it ought to make frequent descents to check its clearness and validity. Kant, as is well known, was inclined to avoid concrete examples on the ground that readers so readily drew the wrong point from them. He was mistaken. Over and over again critics have thrown his imposing machinery out of gear by tossing in some simple case – the saving of a life by a lie, for example – with which the machine, for all its complexity, could not deal. Nelson seems to have a similar feeling; the reasoning moves like a Juggernaut down the line of abstract deductions. After a while one longs for instances, and sometimes feels that the instance that would fully clarify might also disclose cracks in the machine.

At two points Nelson would need, if he were writing today, to fortify his position at the outset. That ethical judgments are objective, that is, true or false, he does not seriously question. To question this successfully would put the whole structure in jeopardy. Well, it *has* been questioned, though not in my judgment successfully, and a present-day moralist could hardly afford to take his case here for granted. Again, Nelson has no doubt that there is synthetic *a priori* knowledge, not only in logic and

mathematics, but also in ethics. I agree, and think the case against it entirely inadequate. But in the present philosophical climate, this opinion is heresy, and demands a competent technical defence if it is to be held at all. That Nelson could and would have offered such a defence is beyond doubt.

Another difficulty has come to the fore in recent years that he might have found it harder to meet, a difficulty that has troubled such British moralists as Ross, Prichard and Carritt. He denies that the moral imperative is hypothetical; duty is not duty because it is a means to anything, however good. The pursuit of goods is the business of culture, not of morality. Now it is surely obvious, as it is to his British colleagues, that duty cannot disregard goods, and Nelson too seems to admit this when he says that in finding our duty, we must take interests into account. There is no 'duty' for ordinary life that may not be suspended if doing it would bring a loss of some great good to ourselves or others. But if so, it is clear that the moral law must be, at least in part, a hypothetical imperative. In short, if duty is to be determined independently of interest, Nelson's position shares the difficulties that entangled Kant. If it may be determined in part or whole by interest, has he not retreated to the teleological ethics he sought to escape?

I do not think such a retreat would be fatal. That rationalism and a teleological ethic can be combined has been strikingly shown by Joseph in his criticism of the English moralists I have mentioned. But Nelson's system is so tightly articulated that a change at any important point would involve a change throughout, and it is idle to conjecture how he might have recast the whole. His belief that ethics can be a science of which reason is the artificer is, I think, right, and in its cool confidence it offers an arresting challenge to current fashions. Whatever the ultimate verdict on his system, he has built an impressive piece of classic architecture, and sometimes these structures continue to shine on their modest hills when current fashions have been forgotten.

BRAND BLANSHARD

Vom Ursprung und Ende der Metaphysik by ERNST TOPITSCH. Springer-Verlag, Vienna, 1958. Pp. 313. DM 26.

This book is an attempt to present a sociological critique of existing metaphysical systems. The author summarizes the results of earlier investigations (among others those of Cassirer and Kelsen) and, on this basis, endeavours to prove that the philosophies of all times and all nations are derived from notions of certain models, particularly the notions of biological development and 'intentional' behaviour. The group of biomorphous models (for instance the mythological notion that the world originated from an egg) is treated very briefly, while the group of 'intentional' models forms the actual basis of Topitsch's investigations. 'Intentional models stem largely from social relationships and systems, from

the family to the state, or else from artistic and technical activity, and must therefore be sub-divided into the sub-groups of socio-morphous and techno-morphous models. With the aid of these models, particular processes or even the entire universe are interpreted as social phenomena or as artifacts. In this way, many apparently "closed" intentional philosophies appear as mere reflections of man's volition and action which are then simply reapplied to the latter; for man's will and activity are to be fitted into the "harmony" of the cosmic society or work of art.

'The present work investigates the different functions of this world-view, and its inherent problems and final solution by modern science and the actual development of modern society. Beyond that, the work attempts to show that at least an integral part of traditional metaphysical and moral philosophizing is rooted in this world-view or else – assuming different values – has developed in opposition to it and can only be understood as a result of this opposition' (p. 3f).

This well-written and readable book will be found invaluable by all sociologists interested in considering philosophical views from this angle. Topitsch's decisive conclusion is probably that scientific thought, in the course of this historic development, has rid itself of 'intentional' ideas, thus making possible the unprejudiced exploration of the world.

Different stages of this development are distinguished. Modern scientific thought began with the emancipation of reason during the Renaissance – man no longer wanted merely to believe, he now wanted to know why things happened. At first this new approach was still steeped in the biomorphous and intentional ideas of ancient and medieval tradition. Copernicus, for instance, justified this heliocentric theory by his assertion that rest was more divine than motion. These intentional ideas were still important for the philosophers of the Renaissance and have even persisted right into modern times. As Karl Loewith showed very clearly in his *Weltgeschichte und Heilsgeschehen*, Hegel, Marx and all the other social philosophers of progress were filled with Messianic hopes of salvation. Topitsch supplies critics of the communist philosophy with weighty arguments – historical materialism is permeated by teleological considerations in disguised form: history is seen as the self-realization of true justice. Even the objectivized values of Marxism are demands rather than economic findings.

The author's attempt to present a sociological critique of philosophy is laudable because it exposes prejudices that have affected scientific thought right into our day. But carried to its logical conclusions Topitsch's approach would lead to a kind of behaviourism. While an analysis of man's external behaviour might well throw light on social and economic problems, it cannot explain man himself. According to Topitsch, once we have shown that philosophies originated from intentional models, we have also proved them to be meaningless. If he were right, the end of metaphysics would at one and the same time spell the end of its subject-matter and of the problem of man's place in the universe and his resulting metaphysical difficulties. But this is by no means the case. As long as

men exist, they will continue to search for a philosophy that gives a meaning to their existence. Once we deny to existence all meaning, simply because we fail to find a meaning in empirical reality, we fall victim to a dangerous nihilism.

The scientist is inclined to forget that his approach is not the only possible one. Restricted as his viewpoint is to the investigation of phenomena, the essential content of the great philosophies must for ever escape him. This becomes obvious from Topitsch's attempt to read intentional models into the beginnings of the Indian and Chinese philosophies. Yet mythical thought, being mostly symbolic thought, defies all such attempts at interpretation. The logical meaning of symbolic thought can only be plumbed with interpretations such as those of C. G. Jung. While logical thought invariably replaced the earlier mythological forms in the course of the philosophic development of all cultures (India, China, Greece, the West), it must be remembered that philosophic thought is basically different from scientific thought, even though both use the same tool: reason. Topitsch completely ignores the development of metaphysics in particular cultures. Thus in discussing Chinese philosophy, he merely mentions the principle of inaction, *wu-wei*, which is realized by man fitting himself into the cosmic harmony (p. 64). However, the *wu-wei* of Chinese philosophy does not mean absolute passivity, but the reasonable activity that is made possible by a realization of the futility of an exaggerated 'instinctuality'. Clearly, if the intellect cannot curb our instinctive urges, there is no room for a moral philosophy that sets limits to our blind desires.

Indian philosophy distinguishes between two types of knowledge: *avidya* and *vidya*. *Avidya* (literally: nescience) refers to purely intellectual knowledge which leads to self-consciousness and self-will. *Vidya* (enlightenment), on the other hand, refers to true philosophic knowledge and wisdom. *Vidya* is not concerned with exploring the external world, but with solving the problems reason itself has created, and with linking rational knowledge to our lives meaningfully.

Topitsch holds that the Chinese notion of submission to the Great Ultimate is as vague as the Stoic principle of conduct in accordance with a universal Logos. All such principles, he claims, are 'commands without content' (p. 64). Accordingly, he looks upon all but empirical judgments as intentional analogies with no claim to scientific validity. Topitsch tries to prove this contention for all philosophies that distinguish between empirical reality and essential reality. When Plato says that man, not God, is responsible for all the ills of this world, he simply means that man's intellect enables him to choose between right and wrong. *Hybris* is what the ancient Greeks called man's exaggerated restlessness, and in *hybris* they saw the source of all evil. They stressed the importance of recognizing what is inessential, i.e. exaggerated activity. *Hybris* is the result of man's ignorance. The search for values transcending empirical reality is not Utopian – these values become apparent once we realize how futile man's activities can be. Human reality, as it is generally understood,

is not the essential reality; it is but the human situation caused by ignorance. If investigations are confined to man's empirical behaviour, they can never arrive at philosophic judgments. While a critique of the great philosophies may well be applied to the contemporary assessments of what is the essential life, it cannot touch their lasting content. To do so would be tantamount to throwing the baby out with the bath water.

Finally, we must examine Topitsch's assertion that value judgments of a socio-cosmic origin entered into Kant's scientific ideal. Kant is said to have tried thereby to defend the universal validity of natural laws, which corresponded with his scientific ideal, from Hume's objections. This assertion rests on the common lack of distinction between empirical knowledge and Kant's transcendental knowledge. Kant fully agreed with Hume as regards empirical knowledge, but while Hume tried to derive causality from experience, the better to protect it against being used for metaphysical speculations, Kant showed that empirical knowledge could not exist without the principle of causality – the ground of all possible experience. He thus achieved two objects: while refuting the claims of speculative metaphysics, he simultaneously asserted the validity of rational concepts that refer exclusively to the objects of experience. On the other hand Topitsch is justified in criticizing Kant's *Critique of Practical Reason* since it does, in fact, contain intentional notions which contradict the analytic findings of the *Critique of Pure Reason*. Thus Kant was the first to overstep the limits he himself had shown to exist for our reason, limits that his successors have again and again transgressed.

One cannot help being struck by the fact that Topitsch has ignored important metaphysical doctrines. A book discussing the beginning and end of metaphysics cannot afford to ignore philosophers such as Bergson, Berkeley, Schelling and Schopenhauer, to mention only a few.

HEINRICH KAESTNER

(Translated from the German)

Raum und Zahl by KURT REIDEMEISTER. Springer-Verlag, Berlin-Goettingen-Heidelberg, 1957. Pp. 151. DM 19.80.

This introduction to mathematical thought concentrates on describing the points of view and methods characteristic of the development of modern mathematical disciplines. By means of examples (taken largely, but not exclusively, from geometry) the book introduces the reader to the modern interpretation of mathematical disciplines as theories about axiomatic structures. Here we cannot give a detailed description of the mathematical contents of the book, important reading though they are for all who wish to gain a rough idea of the modern approach to mathematics. Instead, we shall merely cite the table of contents: I. The Origin of Mathematical Thought. II. Mechanisms. III. Analytic Geometry. IV. Distinction between regions in space: 1. The problem of the orientation of straight lines and the distance between points; 2. The problem of

the orientation of planes and the distance between points; 3. Logic and Ontology. V. Intuition and Concept: 1. The Paradox of Intuition; 2. Ordinal Numbers; 3. Combinatorial Topology. VI. Geometry and Logic. VII. A Foundation for Infinitesimal Calculus. VIII. Carl Friedrich Gauss. IX. Geometry and the Theory of Numbers. X. Prolegomena to a Critical Philosophy.

It is characteristic of the author's philosophic approach that he stresses the importance of analytic thought and reason and subjects even the basic assumptions of a particular viewpoint to a fundamental critique, sparing neither neo-positivism nor existentialism. He calls this the position of critical reason, but, though we must agree with him in that respect, we cannot follow him all the way and, particularly, we must consider his approach to be incomplete when it comes to the role of intuition in mathematics.

Any clarification of the concept of number, he claims, is contingent on the possibility of an axiomatic definition of numbers in which the concept of a field of elements precedes the general concept of number. Certainly this is the case for real numbers, complex numbers, hypercomplex numbers, etc. Natural numbers also can be described by axiomatic structures. However, more recent investigations into the foundations of mathematics, and particularly those arising from Hilbert's proof theory, have shown that, in order to derive logical conclusions from an axiom system, an intuitive grasp of natural number must already have been assumed in so-called meta-mathematics. Parallel attempts are the efforts of the intuitionist school to clarify the intuitive concept of natural number which, for its part, must now be assumed to be essentially *a priori* in Kant's sense.

The problem is more complicated for the case of geometry. The author's belief that the content of those geometric theories which have a logical character is purely inferential, i.e. that it consists of different structures, is certainly true of the construction of abstract geometries. Real space must be considered as the relation between perceptual and physical space. The author is perfectly justified to hold that perceptual space is not *a priori* Euclidean, and that the propositions of geometry are not apodictic (or more precisely, not all its propositions are apodictic), so that the Kantian assumption that Euclidean geometry is a purely intuitive geometry is no longer tenable. Real geometry is largely determined by physics, but physical geometry, in turn, must fulfil the condition that, however unaccustomed we are to it on the sub-microscopic scale, on the macroscopic scale it must correspond with perceptual space, or else perceptual space must unequivocally refer to physical geometry. Now, it cannot be denied that, for instance in visual space, certain geometric facts, e.g. those of a simple topological nature, strike us as irrefutable, and that even the geometry of physics is bound to take these facts into account. These facts, by the way, are not contradicted by relativity, or by any other physical theory; it must, however, be remembered that every attempt to define these geometric facts involves certain abstractions which may well go beyond the purely intuitive content. Thus, if we call Jordan's

curve theorem such a fact, we must remember that, in so doing, we have assumed the continuity of space in Dedekind's sense, which can never be intuitively grasped, since intuition applies only to what is finite and perceivable. Such abstractions are even represented by the concepts 'point', and 'straight line'. On the other hand, we gain the impression that, for instance, something holds for our visual space which corresponds to Jordan's theorem, no matter how difficult we might find it to express this fact without going beyond the sphere of pure intuition. Perhaps this is impossible, but no matter what non-intuitive abstractions we may use, we must somehow account for what is intuitively given.

WILHELM ACKERMANN

(Translated from the German)